UNIVERSITY
OF TASMANIA

Genetic and molecular analysis of two new loci controlling flowering in garden pea

By

A S M Mainul Hasan

School of Natural Sciences

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

University of Tasmania, July 2018

Declaration of originality

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledge in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

Authority of access

This thesis may be made available for loan. Copying and communication of any part of this thesis is prohibited for two years from the date this statement was signed; after that time limited copying and communication is permitted in accordance with the Copyright Act 1968.

_____ Date: 6-07-2018

A S M Mainul Hasan

Abstract

Flowering is one of the key developmental process associated with the life cycle of plant and it is regulated by different environmental factors and endogenous cues. In the model species *Arabidopsis thaliana* a mobile protein, FLOWERING LOCUS T (FT) plays central role to mediate flowering time and expression of *FT* is regulated by photoperiod. While flowering mechanisms are well-understood in *A. thaliana*, knowledge about this process is limited in legume (family *Fabaceae*) which are the second major group of crops after cereals in satisfying the global demand for food and fodder. Due to its short generation time, ability to reproduce via self or cross-pollination and availability of diverse lines, garden pea (*Pisum sativum*) serves as a model legume species for flowering time studies.

Isolation and characterization of mutants have been a key research strategy in order to identify genes responsible for flowering in pea. The current study involved investigation of two novel EMS mutants namely *late3* and *late4* in the background of a cultivated pea line NGB5839 which are extremely late flowering indicating that *LATE3* and *LATE4* are essential for normal promotion of flowering in pea. Detailed phenotypic characterization carried out in the present study showed that both the loci positively regulate various vegetative, reproductive and yield related traits across different growth stages strongly suggesting potential global regulatory role for the underlying genes.

Prior to the present study, synteny between pea and a closely related species, *M. truncatula* was exploited in order to roughly map *LATE3* and *LATE4* at the middle of pea *LGIII* and bottom of *LGV* respectively. These positions were significantly refined in the present study as *LATE3* and *LATE4* were mapped within a narrow genetic intervals in the syntenic regions of *M. truncatula* consisted of 62 (chromosome 3) and 54 (chromosome 7) genes respectively.

Usage of the high-throughput RNA sequencing technology along with phylogenetic, co-segregation analysis and direct sequencing assisted in determining *LATE3* and *LATE4* genes as the pea homologue of *Cyclin Dependent Kinase 8* (*PsCDK8*) and *Cyclin C* (*PsCYCC1*). CDK8 and CYCC1 are components of highly conserved CDK8 module (along with MED12 and MED13) of eukaryotic mediator complex that negatively regulate transcription of various genes involved in different biological processes. Analysis of the entire mediator complex revealed that this

complex is highly conserved in *A. thaliana, M. truncatula* and pea. Phenotypic characterization of loss-of-function mutant of novel flowering gene *CYCC1* in *A. thaliana* exhibited delayed flowering similar to already known similar mutants of *AtCDK8*, *AtMED12* and *AtMED13* which was consistent with the notion that these four genes act together to carry out the same regulatory process.

Genetic interaction and yeast two hybrid assay unveiled complementarity and strong physical interaction between *LATE3* and *LATE4* which strongly suggested that these two genes act in the same regulatory process in a likely inter-dependent manner. Further genetic and regulatory interaction studies involving already known key pea flowering loci showed that *LATE3* and *LATE4* mediate flowering by regulating expression of known genes such as *FTa1*, *FTc*, *PIM* (*PsAP1*), *VEG1*(*PsFULc*), *UNI* (*PsLFY*)and *VEG2* (*PsFD*) and *LF* (*PsTF1c*).

The present study also undertook a systems biology approach for generating relevant hypothesis about the function and regulation of *LATE3* and *LATE4* genes. To this end, STRING v10.5 predicted the potential interactome network of AtCDK8 and AtCYCC1 consisted of members from cell cycle and mediator complex indicating a conserved role for them. In addition, predicted functions of *MtCDK8* and *MtCYCC1* genes obtained from AraNET v2.0 database was similar to that of *PsCDK8*, *PsCYCC1*, *AtCYCC1* genes, thus giving hints of potential conservation of the function of both the genes in plant system. Moreover, PlantTFDB v4.0 database was used to forecast the transcription factors that may regulate the function of *CDK8* and *CYCC1* genes in *M. truncatula* and pea.

As various mediator complex genes in *A. thaliana* are known to be involved in the regulatory processes driving response to different environmental factors, therefore relevant experimentation have shown that both *LATE3* and *LATE4* genes are important for controlling response to general light/darkness, ambient temperature, UVB, heat, mechanical wound and salt stress in pea.

Overall, the present study provided significant understanding about the role of *LATE3*/*PsCDK8* and *LATE4*/*PsCYCC1* in mediating diverse range of reproductive, developmental and adaptive characteristics in pea where flowering time was the most important trait.

Acknowledgements

During my journey as a PhD student at the School of Natural Sciences, University of Tasmania I have received assistance from many individuals and I would like to humbly convey my gratitude to all of them. First of all, I would like to thank my primary supervisor and mentor Dr. Jim Weller who always managed time for me whenever I sought guidance for any aspects of the project be it laboratory/glasshouse work, writing the thesis or attending conferences. I indeed believe that I have become a better scientist over the past three and half years where his leadership played a significant role. Likewise, I am thankful to my second supervisor Dr. Valerie Hecht whose technical expertise and in-depth knowledge about various aspects of flowering time regulation have contributed immensely towards my development as a PhD student.

Special thanks goes to my colleague Jackie Vander Schoor who has played significant role in the development of this project at different stages. Another person who has offered guidance and inspiration regarding my academic responsibilities and ownership of my PhD project is my graduate research coordinator Dr. Erik Wapstra and therefore I am beholden to him.

Next, I would like to acknowledge helps that I have received from my colleague Raul Ortega for useful discussion about various issues related to my PhD project. I am also conveying similar humility to Dr. Lu Wang specifically for the *A. thaliana* part of the project. In addition, I am grateful for various suggestions and scientific brainstorming by my former and current group mates Dr. Frances Sussmilch, Dr. Steven Ridge, Dr. Vinodan Rajendran, Owen Williams, Beatriz Contreras, Andy Rubenach and Kelsey Picard.

Colleagues in the glass house Michelle Lang and Tracy Winterbottom were fantastic and cooperative and their help from time to time is highly appreciated. The same humbleness goes to recognize the assistance of colleagues from CSL molecular biology laboratory Adam Smolenski and Sharee McCammon. I would also like to thank the bigger family of the School of Natural Sciences specially those with whom I worked as part of the school's student organization in organizing various social events in the first year of the PhD. In addition, I would like to thank University of Tasmania and Australian Research council for providing financial

support for this project. Moreover, I am thankful to Australian Society of Plant Scientists (ASPS) for providing grant in order to be able to attend the Combio 2017 conference in Adelaide.

I would like to specifically mention the contributions being made by my previous professional mentors Prof. Dr. Dr. Wolfgang Friedt, late Dr. Patrick Schweizer and Dr. Thorsten Schnurbusch towards my development which enabled me to undertake intensive plant molecular genetics based research in the form of a PhD.

I am highly indebted to my lovely wife who constantly provided support and took care of various aspects of my life during this period. At this specific junction of my life, I would also like to convey utmost respect and gratitude to my mother, late father and elder brothers without whom I would not have been where I am today. Besides, I would like to thank my fellow compatriots living in Hobart for their consistent assistance and care about different matters during my stay as a PhD student in this city.

Conference proceedings from this project

Invited talk: A.S.M Mainul Hasan, Valerie Hecht, Jacqueline K. Vander Schoor, Jim Weller, October 30-November 1, 2017. Genetic and molecular analysis of two new loci controlling flowering in garden pea, *Pisum sativum*. In: Australasian Genomics Technologies Association Conference, Hobart, Australia.

Poster presentation: A.S.M Mainul Hasan, Valerie Hecht, Jacqueline K. Vander Schoor, Jim Weller, October 4-6, 2017. Genetic and molecular analysis of two new loci controlling flowering in garden pea, *Pisum sativum*. In: Combio 2017, Adelaide Convention Centre, Adelaide, Australia.

Abbreviation

3AT	3-amino-1,2,4-triazole
ABA	Abscisic acid
AFR	*SAP30 FUNCTION-RELATED*
AGL18	*AGAMOUS LIKE 18*
ANOVA	Analysis of variance
AP1	*APETALA1*
bp	Base pair
bZIP	Basic leucine zipper
CAPS	Cleaved amplified polymorphic sequence
CDF	*CYCLING DOF FACTOR*
CDK8	*Cyclin Dependent Kinase 8*
cDNA	Complemetary Deoxyribonucleic Acid
cM	Centi Morgan
CO	CONSTANS
Col-0	Columbia 0
CRY2	*CRYPTOCHROME 2*
CSM	Cryptic splice motif
CSS	Cryptic splice site
cv.	Cultivar
CYCC1	*Cyclin C*
dCAPS	Derived cleaved amplified polymorphic sequence
DET	*DETERMINATE*
DNA	Deoxyribonucleic acid
DNE	*DAY NEUTRAL*
EFS	*EARLY FLOWERING IN SHORT DAYS*
ELF3	*EARLY FLOWERING 3*
ELF4	*EARLY FLOWERING 4*
EMS	Ethyl methane sulfonate
EST	Expressed sequence tag
EZA1	*ENHANCER OF ZESTE 1*
FKF1	*FLAVIN KELCH BOX 1*
FLC	*FLOWERING LOCUS C*
FM	Floral meristem
FRI	*FRIGIDA*
FT	*FLOWERING LOCUS T*
FUL	*FRUITFULL*
GA	Gibberellic acid
gDNA	Genomic Deoxyribonucleic Acid
GI	*GIGANTEA*
GID	*GIBBERELLIC ACID-INSENSITIVE DWARF*

GTF	General transcription factors
HAT	Histone acetylase
HDAC	Histone Deacetylase Complex
HR	*HIGH RESPONSE TO PHOTOPERIOD*
HRM	High resolution melt
HSP	Heat shock protein
I1M	Primary inflorescence meristem
I2M	Secondary inflorescence meristem
IRN	Initial reproductive nodes
IRVN	Inflorescence reversed vegetative nodes
JA	Jasmonic acid
JMJ14	Jumonji14
Kb	Kilo base pair
LATE1	*LATE BLOOMER 1*
LATE2	*LATE BLOOMER 2*
LATE3	*LATE BLOOMER 3*
LATE4	*LATE BLOOMER 4*
LD	Long day
LF	*LATE FLOWERING*
LFY	*LEAFY*
LG	Linkage group
LHP1	*LIKE HETEROCHROMATIN 1*
lncRNA	Long non coding RNA
LRN	Later reproductive nodes
LUX	*LUX ARRHYTHMO*
Mb	Mega base pairs
MED12	*MEDIATOR 12*
MED13	*MEDIATOR 13*
NFD	Node of flower development
NFI	Node of flower intiation
NGS	Next generation sequencing
NS	Non significant
NTC	Non template control
ORF	Open reading frame
PAF1c	RNA Polymerase II Associated Factor 1 complex
PCR	Polymerase chain reaction
PEBP	Phosphatidylethanolamine-binding protein
PHYA	*PHYTOCHROME A*
PHYB	*PHYTOCHROME B*
PIM	*PROLIFERATING INFLORESCENCE MERISTEM*
Pol II	RNA Polymerase II
PPD	PHOTOPERIOD RESPONSE
PRC1	Polycomb repressive complex 1
PRC2	Polycomb repressive complex 2

PTC	Premature stop codon
PUF4	*PHYTOCHROME INTERACTING FACTOR4*
Q-RT-PCR	Quantitative reverse transcriptase polymerase chain reaction
REF6	*RELATIVE OF EARLY FLOWERING 6*
RN	Reproductive nodes
RNA	Ribonucleic acid
SAM	Shoot apical meristem
SAS	Shade avoidance syndrome
SC	Synthetic complete medium
SD	Short day
SDG	*SET domain gene*
SDW	Sterile distilled water
SE	Standard error
SN	STERILE NODE
SNP	Single nucleotide polymorphism
SOC1	*SUPRESSOR OF OVEREXPRESSION OF CONSTANS 1*
SS	Splice site
SUF4	*SUPPRESSOR of FRI 4*
SVP	*SHORT VEGETATIVE PHASE*
SWN	*SWINGER*
T6P	Trehalose-6-phosphate
TF	Transcription factor
TFIIa	*TRANSCRIPTION FACTOR IIa*
TFL1	*TERMINAL FLOWER 1*
Tm	Optimal annealing tempearature
TN	Total nodes
TXR	*TRITHORAX RELATED*
UNI	*UNIFOLIATA*
uORF	Upstream open reading frame
UTR	Untranslated region
UV	Ultra violet
VEG1	*VEGETATIVE 1*
VEG2	*VEGETATIVE 2*
VM	Vegetative meristem
VRN1	*VERNALIZATION 1*
WGS	Whole genome sequencing
WT	Wild type

Table of contents

Chapter 1: General introduction

1.1 General biology of flowering

During the life cycle of a plant, it is very important for the plant to choose the right time and condition to flower. The onset of flowering is a key developmental transition in the plant life cycle and is regulated by different environmental factors such as photoperiod, temperature and endogenous cues like circadian clock, gibberellic acid and developmental age. Since flowering time is related to the ultimate yield in plants, therefore it is of great interest for plant scientists to understand the mechanism underlying this very complex trait.

The plant organ on which flowers are produced is known as inflorescence. During the initial vegetative stage of plant growth, the shoot apical meristem (SAM) develops leaves and shoots, but at a certain point it transforms into an inflorescence meristem once floral transition has occurred in the plant (Benlloch et al. 2007). Based on branching pattern and position of flowers, different types of inflorescence architecture have evolved among plants.

Day length or photoperiod is one of the most important environmental factors determining flowering in plants, and many plants require exposure to specific photoperiods for induction of flowering. Plants can be divided into three broad categories based on their photoperiod responses (Andres and Coupland 2012). Long-day (LD) plants flower when length of the day is above a particular threshold, whereas short-day (SD) plants flower when the length of night is above a particular threshold. Day-neutral plants flower irrespective of the length of day and night.

Responsiveness to photoperiod is dependent on an endogenous timekeeping mechanism called circadian clock in the plants. The word circadian is derived from the Latin *circa* (about) and *diem* (a day). The circadian system has three main components: a central and self-sustaining clock, various input pathways that adjust clock function in accordance to environmental signals, and output pathways that mediate clock regulation of various cellular processes and modulate key features of plant development (Hsu and Harmer 2014). Temperature is another important environmental factor that has a profound effect on flowering mechanism in many plant species (Andres and Coupland 2012). Many plants from temperate regions (e.g. winter habit genotypes of barley/wheat and pea) need to experience a period of cold during winter before being able to flower in the following summer, a process

known as vernalization. In contrast, there are other plant species from warmer regions (e.g. rice and soybean) that are insensitive to vernalization and can flower without exposure to low temperature.

Arabidopsis thaliana has been extensively used as a model system in order to study different regulatory pathways in plants. *A. thaliana* has been preferred by plant scientists all over the world for understanding physiology and molecular biology of flowering for several reasons: its small genome size (~135 Mbp), easy handling, short life cycle, comparatively fast generation time and prolific seed production (Analysis of the genome sequence of the flowering plant Arabidopsis thaliana 2000). Most of the current knowledge on flowering is based on mutant analysis in *A. thaliana*.

1.2 Different factors and regulatory pathways mediating flowering time in *Arabidopsis thaliana*

1.2.1 Flowering locus T (FT) as a central component

Extensive studies in *Arabidopsis thaliana,* a facultative LD plant*,* have provided a detailed understanding of the genes and genetic interactions underlying the transition to flowering and its regulation by photoperiod and temperature. Flowering in *A. thaliana* is regulated by a key mobile protein called FLOWERING LOCUS T (FT), a member of the phosphatidylethanolamine-binding protein (PEBP) family (Pin and Nilsson 2012). FT is also known as florigen and acts as a key integrator for the signals generated from different flowering regulatory pathways (Figure 1.1) (Andres and Coupland 2012; Song et al. 2013; Johansson and Staiger 2015; Song et al. 2015). Upon receiving different endogenous and exogenous signals in leaves, the *FT* gene is expressed and the FT protein moves to the SAM through phloem where it activates meristem identity genes to induce flowering (Andres and Coupland 2012; Song et al. 2015).

After transcriptional induction of *FT* in the vasculature of the leaf, the FT protein moves via companion cells of the phloem to the SAM where it forms a florigen activation complex by interacting with a basic leucine zipper (bZIP) domain transcription factor FD and 14-3-3 protein (Abe et al. 2005; Jang et al. 2009; Wigge et al. 2005; Golembeski and Imaizumi 2015). The FT-FD interaction activates the transcription of several MADS box transcription factors at the SAM, including the integrator gene *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1*

(*SOC1*) (Borner et al. 2000; Samach et al. 2000), and *APETALA1* (*AP1*) which has a dual role in floral induction and in floral meristem identity (Wigge et al. 2005). The unrelated protein LEAFY (LFY) is also indirectly induced by FT-FD and acts with AP1 to initiate flower formation (Andres and Coupland 2012). Other transcription factors such as AGAMOUS-LIKE-24 (AGL24), members of the SQUAMOSA BINDING PROTEIN LIKE (SPL) family and a MADS box FRUITFULL (FUL) act as positive regulators of *AP1* and *LFY*. TERMINAL FLOWER 1 (TFL1) is a protein similar to FT and functions in negatively regulating expression of *AP1* and *LFY* in the SAM. In young floral primordium, *AP1* and *LFY* suppress the function of *TFL1* (Liljegren et al. 1999; Shannon and Meeks-Wagner 1993; Andres and Coupland 2012; Johansson and Staiger 2015). Repressive activity of *TFL1* on its targets is dependent on *FD*. Thus, FD plays dual role in floral initiation based on its interaction with FT or TFL1 (Hanano and Goto 2011).

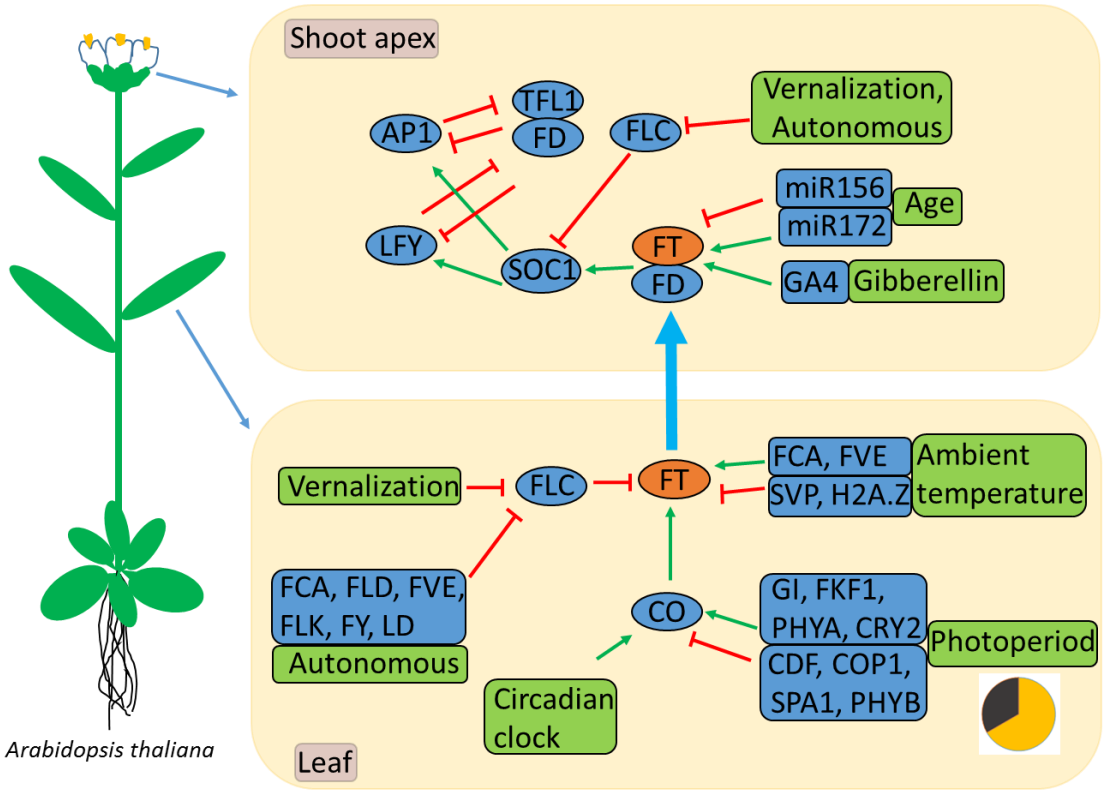


Figure 1.1. Regulation of flowering time in *A. thaliana* through various pathways. Green arrows mean gene promotion whereas red perpendicular lines mean gene repression. Full names of the genes mentioned in text.

1.2.2 Light and clock regulated photoperiodic pathway

In *A. thaliana* grown under LD condition, *FT* expression peaks during later part of the day, as a result of direct activation by the *CONSTANS* (*CO*) gene (Figure 1.1) (Suarez-Lopez et al. 2001; Tiwari et al. 2010; Song et al. 2012b; Johansson and Staiger 2015; Song et al. 2015). A repressor known as *CYCLING DOF FACTOR* (*CDF*) negatively regulates *CO* transcription during the morning. This repression is overcome in the afternoon of a long day when a circadian clock protein GIGANTEA (GI) interacts with a blue light photoreceptor FLAVIN KELCH BOX 1 (FKF1) to degrade CDF (Sawa et al. 2007). At the post-translational level, CO is ubiquitinylated at night by a complex between CONSTITUTIVE PHOTOMORPHOGENIC 1 (COP1) and SUPPRESSOR OF PHYTOCHROME A (SPA1) (Jang et al. 2008; Laubinger et al. 2006; Liu et al. 2008). The SPA1-COP1 complex is unable to function under light and is repressed by a far-red light receptor PHYTOCHROME A (PHYA) and blue-light receptor CRYPTOCHROME 2 (CRY2) in the afternoon of a long day (Briggs and Olney 2001; Zuo et al. 2011). This results in stabilization of CO protein and significant up-regulation of *FT* during this period.

The rhythmic expression of certain key genes, particularly *CO*, is central to the photoperiod response mechanism in *A. thaliana*, and the circadian clock plays important role in establishing these rhythms and thereby regulating flowering time (Figure 1.1) (Hsu and Harmer 2014; Shim et al. 2017). The circadian clock is currently considered to consist of a network of genes whose expression is controlled in a number of interlocking feedback loops. Clock genes are often categorized according to the time of day when their expression is highest (Hsu and Harmer 2014; Johansson and Staiger 2015; Shim and Imaizumi 2015). Important morning-phased components include MYB like transcription factors CIRCADIAN CLOCK-ASSOCIATED (CCA1) and LATE ELONGATED HYPOCOTYL (LHY). Evening phased components include TFs TIMING OF CAB EXPRESSION 1 (TOC1), PRR5, CCA1 HIKING EXPEDITION (CHE), LUX ARRHYTHMO (LUX), BROTHER OF LUX ARRHYTHMO (BOA), and transcriptional regulators EARLY FLOWERING 3 (ELF3) and ELF4. In between are day phased components such as PSEUDO-RESPONSE REGULATOR (PRR9), PRR7 and the nuclear proteins NIGHT LIGHT-INDUCIBLE and CLOCK REGULATED GENE 1 (LNK1) and LNK2, and the afternoon-phased transcription factors REVEILLE 8 (RVE8), RVE4 and RVE 6. Morning phased genes *CCA1* and *LHY* repress the expression of evening phased genes *TOC1*, *LUX*, *ELF3*, *ELF4* (Hsu and Harmer 2014; Shim et al. 2017). The expression of *CCA1* and *LHY* is in turn suppressed by day

phased genes *PRR9*, *PRR7*, *LNK1* and *LNK2*. Afternoon phased genes *RVE8*, *RVE4* and *RVE6* upregulate expression of evening genes. In contrast, evening phased genes negatively control function of the morning and day phased genes.

Several key flowering genes in the photoperiod pathway including *CO*, *GI*, *FKF1* and *CDF* show clock-driven rhythmic expression. The clock regulates *CO* expression and thus it acts as a key integrator gene between the circadian clock and photoperiodic pathway of flowering (Suarez-Lopez et al. 2001). *CCA1* and *LHY* upregulate *CDF* genes early in the day (Shim and Imaizumi 2015). In contrast, *CDF* is repressed by *PRR9*, *PRR7* and *PRR5* during the afternoon. *CCA1* and *LHY* also suppress function of *GI* in the morning (Alabadi et al. 2001) while *TOC1*/*PRR1* does so in the evening (Huang et al. 2012; Shim and Imaizumi 2015). In the afternoon, expression of *GI* and *FKF1* is induced by *RVE8* (Rugnone et al. 2013). LNK1 and LNK2 form a complex with RVE8 in vivo, thereby assisting activity of RVE8 (Xie et al. 2014). Both *GI* and *FKF1* are needed for activation of *CO*. Thus, the clock ensures significant rise in the expression of *CO*, which ultimately results in higher abundance of FT protein before dusk.

1.2.3 Vernalization pathway

Many plant species from temperate regions need exposure to certain period of cold before inception of flowering. This process is called vernalization and it ensures that the plant undergoes only vegetative growth during the low temperature of winter and is then able to flower in the following spring when more favourable temperature conditions return. In *A. thaliana*, a MADS-box transcription factor FLOWERING LOCUS C (FLC) plays a central role in repressing flowering in the absence of cold, acting to suppress the expression of *FT* in leaf and of *FD* and *SOC1* in the SAM (Figure 1.1) (He 2012; Ream et al. 2012; Whittaker and Dean 2017a). Various chromatin modifiers plays important role to regulate expression of *FLC* (Crevillen and Dean 2011; Jiang et al. 2008; Kim et al. 2009; He 2012).

1.2.3.1 Promotion of FLC function via different chromatin modifiers

FRIGIDA (FRI) is a gene that encodes a plant specific scaffold protein and it plays a crucial role in enhancing function of *FLC* and thereby repressing flowering in *A. thaliana*. Various chromatin modifiers and plant specific components are known to control *FLC* function in *A. thaliana* lines carrying functional *FRI* allele (Crevillen and Dean 2011). FRI forms a putative transcription activator complex, FRIc in combination with two plant-specific factors namely

FRIGIDA like 1 (FRL1) and FRIGIDA-ESSENTIAL 1 (FES1) and two components called SUPPRESSOR of FRI 4 (SUF4) and FLC EXPRESSOR (FLX) (Figure 1.2) (Andersson et al. 2008; Schmitz et al. 2005; Kim et al. 2006).

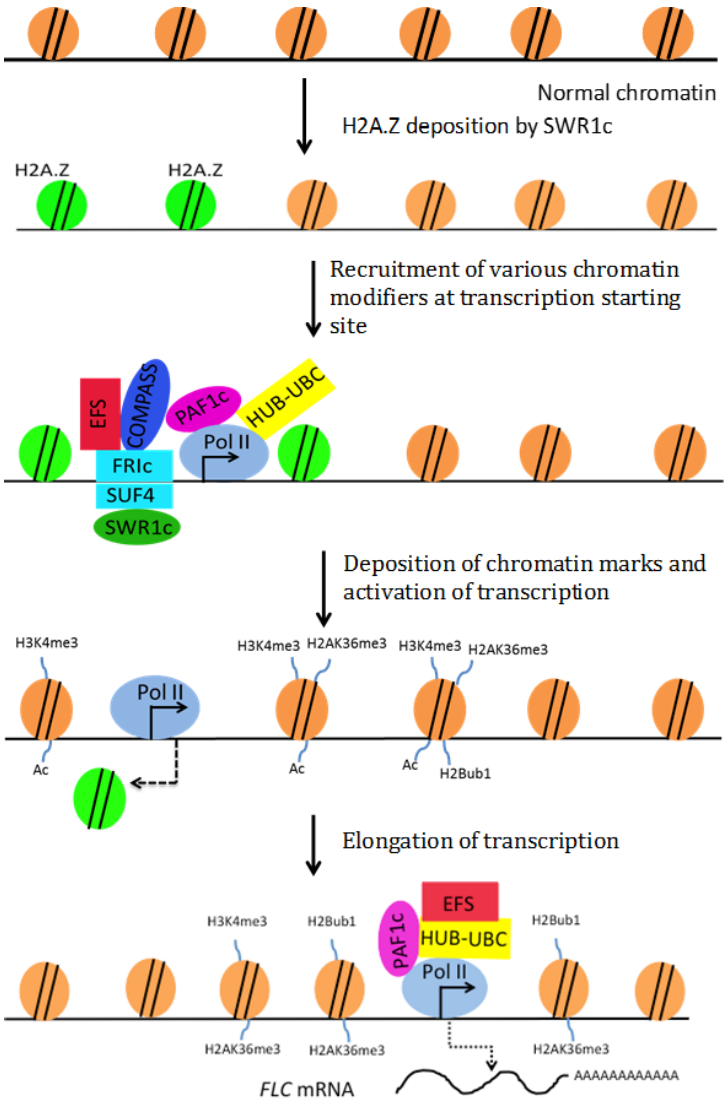


Figure 1.2. FRI mediated activation of FLC through different chromatin modifiers. Full names of the genes mentioned in text. Figure adapted from He 2012 and Zografou T. and Turck F., 2013.

SWR1c, an ATPase chromatin-remodeling complex facilitates *FLC* activation by substituting nucleosomes carrying canonical H2A with a histone variant H2A.Z at the 5' UTR of *FLC* (Figure 1.2) (Choi et al. 2007; Deal et al. 2007). Besides, RNA Polymerase II Associated Factor 1 complex (PAF1c), a chromatin modifier acts as a site of recruitment for histone modifying enzymes at the time of transcription initiation and elongation ultimately resulting in

deposition of H3K4me3, H3K36me2/me3 marks on *FLC* chromatin, which then switch on *FLC* expression (He et al. 2004; Xu et al. 2008). Likewise, H3K4me3 is deposited at the transcription start sites (TSS) of *FLC* by COMPASS- like H3K4 methyltransferase complexes (Jiang et al. 2011). In addition, flowering is delayed by *Early Flowering in Short Days (EFS)* which elevates *FLC* expression through accumulation of H3K36me2/me3 across *FLC* locus (Ko et al. 2010; Xu et al. 2008) and also by HUB-UBC (H2Bub1 complex) that assembles genome-wide H2Bub1 specially at *FLC* (Figure 1.2) (Cao et al. 2008; Gu et al. 2009; Xu et al. 2008).

1.2.3.2 FLC repression during vernalization and post-vernalization winter memory

In the winter annual accessions of *A. thaliana*, *FLC* is repressed during the winter and this repressive state is maintained during the following spring upon return to warm temperature through memorization of cold exposure, which is mitotically stable. This process involves activity of long non-coding RNAs (lncRNAs) and polycomb group proteins (PcG) (Dennis and Peacock 2007; Kim et al. 2009; Angel et al. 2011; He 2012; Zografou and Turck 2013; Whittaker and Dean 2017b).

Upon exposure to cold temperature during winter, two different types of lncRNAs namely COOLAIR and COLDAIR are transcribed from the FLC transcript in *A. thaliana* causing its repression. Among these, COOLAIR is transcribed from the 3' UTR of *FLC* in an antisense direction whereas COLDAIR is transcribed from the first intron of *FLC* in the sense direction (Heo and Sung 2011; Ietswaart et al. ; Swiezewski et al. 2009). The specific role of *COOLAIR* in repressing *FLC* is not fully understood. On contrary, *COLDAIR* plays role in this regard by recruiting polycomb repressive complex 2 (PRC2) that initiates H3K27me3 formation at *FLC* nucleation site (Figure 1.3) (Heo and Sung 2011). Together with *VERNALIZATION INSENSITIVE 3* (*VIN3*) that encodes a plant homeodomain protein, PRC2 forms a PHD-PRC2 complex during vernalization which suppress *FLC* expression by deposition of H3K27me3 repressive marks (He 2012).

The polycomb repressive complex 1 (PRC1) maintains *FLC* repression upon return to warm temperature (Wood et al. 2006; De Lucia et al. 2008; He 2012; Zografou and Turck 2013). PRC1 component, LIKE HETEROCHROMATIN 1 (LHP1) recognizes H3K27me3 mark and spread this mark along the entire *FLC* locus after vernalization (Fig. 1.3) (Angel et al. 2011; Finnegan and Dennis 2007). VERNALIZATION 1 (VRN1) functions in cooperation with LHP1 in this

process (Turck et al. 2007; Zheng and Chen 2011). Besides, EMBRYONIC FLOWER 1 (EMF1) functions in silencing *FLC* by depositing H2AK119ub at *FLC* (Figure 1.3) (Bratzel et al. 2010). Thus, *FLC* is locked at a repressed state and it is maintained stably via cell division upon temperature elevation after winter. As a result, flowering is initiated during spring.

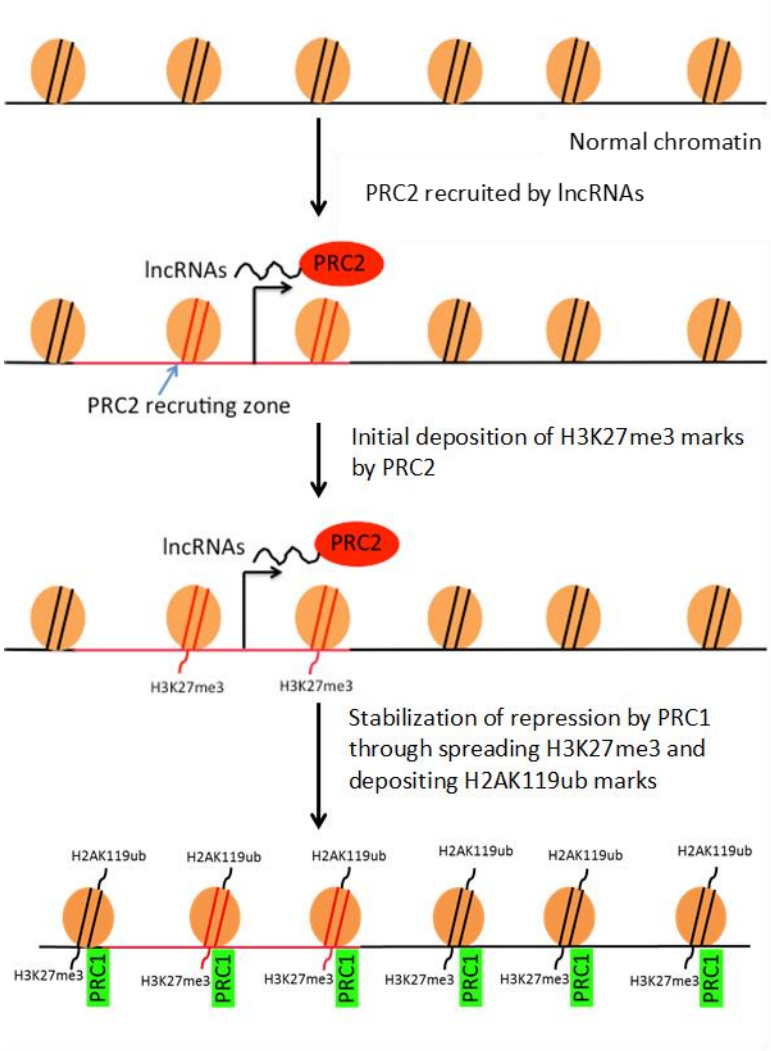


Figure 1.3. Repression of FLC through different chromatin modifiers. Full names of the genes mentioned in text. Figure adapted from Zografou T. and Turck F., 2013.

1.2.4 Autonomous/Constitutive pathway

Components of the autonomous/constitutive pathway play positive role in flowering regardless of day length by suppressing activity of the major floral repressor gene *FLC* at chromatin, DNA, RNA and protein level (Cheng et al. 2017). In *A thaliana*, a protein complex consisted of four histone deacetylases such as FLD, FVE, HDA5 and HDA6 regulate H3 or H4

deacetylation of *FLC* chromatin leading to repression of its expression (He et al. 2003; Domagalska et al. 2007; Yu et al. 2011a; Luo et al. 2015). Besides, two histone demethylases LD and REF6 suppress *FLC* expression by altering H3K4 demethylation status of *FLC* chromatin.

Plant specific RNA processing proteins FCA and FPA act in mediating alternative polyadenylation and 3'end formation of *FLC* (Schomburg et al. 2001; Liu et al. 2007; Hornyik et al. 2010; Liu and Mara 2010). Two other 3' end processing proteins namely FY and PCFS4 have known role for interaction with FCA (Simpson et al. 2003; Xing et al. 2008) and thus all these four proteins may form a complex to regulate expression of *FLC* (Cheng et al. 2017). Another gene *FLD* encodes histone demethylase interacts genetically with *FCA* and *FPA* and thereby probably links chromatin regulation with RNA processing of *FLC* (Liu et al. 2007).

Post-translation modification of FLC protein is also carried out by the autonomous pathway. In this case, casein kinase II (CK2) and protein phosphatase 2A (PP2A) repress FLC by phosphorylation and dephosphorylation (Mulekar and Huq 2015; Mulekar et al. 2012).

1.2.5 Ambient temperature regulated thermosensory pathway

Distinct from vernalization response, flowering in *A. thaliana* is also responsive to differences in ambient growth temperature (Figure 1.1). Many of the genes of the autonomous pathway are involved in this process (Capovilla et al. 2015). Relative to normal growth temperatures, flowering is delayed under LD at 16° C due to repression of *FT* imposed by the floral repressor gene *SHORT VEGETATIVE PHASE* (*SVP*) (Lee et al. 2007). In contrast, *FVE* and *FCA* genes promote *FT* expression by suppressing *SVP* at 23° C resulting in early flowering. Under SD conditions, a repressor H2A.Z binds to the promoter region of *FT* at low temperature slowing down flowering initiation process. H2A.Z is the most evolutionarily conserved histone variant of the H2A histone family and in comparison to regular nucleosomes, causes more compact binding of DNA, resulting in reduced access of transcription factors and lower activity of RNA polymerase II (Kumar and Wigge 2010; van Daal et al. 1990). It has recently been proposed that H2A.Z acts as a sensor of changes in ambient temperature. Repression of *FT* imposed by H2A.Z is overcome when H2A.Z is removed from *FT* promoter due to elevation in temperature. As a result, a basic helix-loop-helix (bHLH) transcription factor, PHYTOCHROME INTERACTING FACTOR 4 (PIF4) is able to bind to the *FT* promoter, activate *FT* transcription, and thus accelerate flowering (Kumar et al. 2012).

1.2.6 Developmental age dependent pathway

Developmental age is another factor that contribute in determining the induction of flowering. As in the case of environmental variables, information about age is ultimately mainly integrated through expression of *FT*, *SPL*, *SOC1*, and meristem identity genes *AP1* and *LFY*. Two key factors in age-dependent flowering control are the miRNAs, e.g., miR156 and miR172, whose expression is dependent on plant age (Figure 1.1) (Huijser and Schmid 2011; Capovilla et al. 2015). miR156 is more highly expressed in young plants than adult ones and targets 11 of the 17 *A. thaliana SPL* genes for down-regulation (Hong and Jackson 2015; Huijser and Schmid 2011). *SPL* genes positively regulate function of the floral integrator and meristem identity genes such as *FT*, *SOC1*, *AP1, LFY*. Thus, miR156 prevents floral formation until the plants reach a certain stage of development. As the plant gets older, miR156 expression is reduced and *SPL*s are released from repression causing upregulation of miR172 and also flowering genes *FT*, *SOC1*, *AP1*, *LFY* (Yamaguchi et al. 2009; Kim et al. 2012). miR172 becomes more highly expressed as the plant becomes more mature, and facilitates flowering by suppressing function of floral repressor genes like *SMZ*, *SNZ*, *TOE1*, *TOE2*, *TOE3* (Aukerman and Sakai 2003; Mathieu et al. 2009).

1.2.7 Gibberellic acid mediated pathway

Gibberellic acid (GA) promotes flowering in *A. thaliana* by acting as a growth enhancer (Figure 1.1) (Fornara et al. 2010). GA regulates *FT* transcription and flowering under LD condition by modulating the function of DELLA proteins. Bioactive GA, GA4 degrade DELLA proteins in the presence of GA receptors such as GIBBERELLIC ACID-INSENSITIVE DWARF 1 (GID1a, GID1b, GID1c) which results in upregulation of *FT* (Nakajima et al. 2006; Daviere and Achard 2013). DELLA proteins are negative regulators of GA signalling and they suppress function of *SPL*s both in leaf (*SPL3*) and shoot apex (*SPL3*, *SPL4*, *SPL5*) (Yu et al. 2012; Galvao et al. 2012). Under LD condition, DELLA delays flowering by acting negatively upon CO and thereby suppressing CO-FT regulated flowering (Galvao et al. 2012; Wang et al. 2016).

GA 20-oxidase (GA20ox) and GA 3-oxidase (GA3ox) regulate several oxidation steps in order to synthesize bioactive GA4 (Fleet and Sun 2005; Osnato et al. 2012). ASYMMETRIC LEAVES 1 (AS1) that is a MYB transcription factor upregulates *GA20ox1* expression (Masaya et al. 2010; Guo et al. 2008b; Song et al. 2012a). AS1 enhances *FT* expression by forming a complex with

CO and binding to *FT* promoter (Song et al. 2012a). So, AS1 promote flowering by regulating GA4 synthesis via GA20ox and also through CO interaction at *FT* locus.

1.2.8 Epigenetic regulation of flowering with relevance to FT expression

In *A. thaliana*, expression of *FT* is regulated by different chromatin modifiers such as SWR1c, PRC2, PRC1, Relative of Early Flowering 6 (REF6) H3K27 demethylase and Jumonji14 (JMJ14) H3K4 demethylase (Figure 1.4) (He 2012).

The PRC2 complex act in accumulating repressive H3K27me3 marks at *FT* locus and thereby prevent its expression in the vasculature (Farrona et al. 2011; Jiang et al. 2008). The H3K27me3 mark is read by the PRC1 component LHP1 which functions in spreading this mark along the *FT* locus (Turck et al. 2007) up to 5kb upstream of transcription start site. Besides, SWR1c complex deposits histone variant H2A.Z at the transcription starting site of *FT* and relative increase in ambient temperature from 17° C to 27° C causes removal of H2A.Z by PIF4 which results in activation of *FT* transcription via Pol II (Kumar and Wigge 2010). Besides, *FT* expression is promoted by REF6, a JmjC-domain H3K27 demethylase as it performs H3K27 demethylation at *FT*.

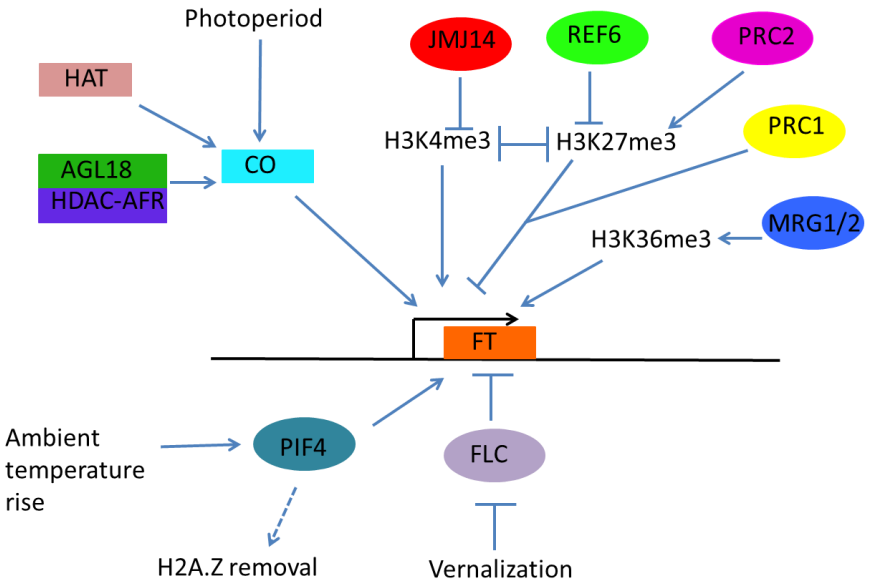


Figure 1.4. *FT* gene regulation through epigenetic mechanism. Arrows mean gene promotion whereas perpendicular lines mean gene repression. Full names of the genes mentioned in text. Dotted arrow means removal. Figure adapted from He 2012.

Recent studies have revealed that *A. thaliana* Morf Related Gene (MRG) Group proteins MRG1 and MRG2 act in reading promotive H3K4me3 and H3K36me3 marks at *FT* chromatin and they also assist in stable CO binding to *FT* promoter region leading to activation of *FT* expression (Bu et al. 2014). The JMJ14 H3K4 demethylase functions in H3K4 demethylation by binding to *FT* chromatin and thereby repress *FT* expression (He 2012). Thus, the relative proportion of repressive H3K27me3 and promotive H3K4me3 marks around the *FT* locus is important for effective repression or expression of *FT*, respectively which then influences flowering accordingly.

Histone acetylation and deacetylation is also crucial in regulating *FT* function. In *A. thaliana*, two functional relatives of the yeast SAP30, known as SAP30 FUNCTION-RELATED 1 (AFR1) and SAP30 FUNCTION-RELATED 2 (AFR2) function as a component of the Histone Decetylase (HDAC) complex (Gu et al. 2013). The HDAC-AFR1 or HDAC-AFR2 complex maintain the acetylation level at *FT* locus upon *CO* mediated activation of *FT* at the end of long day (LD). Besides, an unknown histone acetylase (HAT) may also be recruited by CO at *FT* during this time.

1.3 Flowering control mechanism in agronomically important crops

Flowering studies and development of genomic resources in crop species has lagged behind *A. thaliana*, but over the last decade much progress has also been made in a number of systems. Studies in various crop species have resulted in identification of many of the *A. thaliana* orthologous flowering genes suggesting evolutionary conservation of these genes and flowering control pathways across species. However, the regulation of flowering in many of these species is not entirely similar to *A. thaliana,* and it is becoming apparent that distinct regulatory components and mechanisms have evolved (Andres and Coupland 2012; Abelenda et al. 2014; Weller and Ortega 2015; Hill and Li 2016).

Cereals are the most agronomically important crops in the world (Hill and Li 2016) and are classified within the family *Poaceae* (grasses). Agriculturally important cereal plants include rice, wheat, barley, maize and sorghum. They are the major food source across the world and provide carbohydrates, proteins and different vitamins to consumers. Therefore, it is of great interest for the global scientific community to gain understanding about the flowering time regulatory mechanism in these plants.

Legumes are the second major crop group after cereals (Weller and Ortega 2015). Legumes (family *Fabaceae*) are divided into three different sub-families namely *Papilionoideae*, *Mimosoideae* and *Caesalpinioideae*. Agronomically important legumes are categorized under two subclades of *Papilionoideae*. These are the "galegoid" clade or cool season legumes consisting of chickpea, pea, faba bean, lentil, clovers, medics and the sister "phaseoloid" clade or warm season legumes comprising of soybean, cowpea, common bean and some other bean species (Cronk et al. 2006). In particular they have a valuable role in many agricultural systems due to their ability to fix atmospheric nitrogen via symbiosis with soil microorganisms (Smýkal et al. 2012). They are an important source of dietary protein in many parts of the world, supplying 1/3 of direct human protein intake globally. They also contribute significantly to satisfy the global demand for animal food and fodder.

Flowering control mechanism in some key SD and LD plants from cereals, legumes and nightshades are mentioned below:

1.3.1 Rice (Oryza sativa)

Rice is a model cereal that grows under SD condition in tropical areas and does not need vernalization for flowering unlike *A. thaliana*. Orthologues of *A. thaliana CO* and *FT*/*TSF* genes have been identified in rice, which are called *HEADING DATE 1* (*Hd1*) and *Hd3A/RFT1* respectively (Hayama et al. 2003; Andres and Coupland 2012). The AtGI-AtCO-AtFT module of photoperiod dependent flowering of *A. thaliana* is highly conserved in rice through similar OsGI-OsHd1-OsHd3a module (Tsuji et al. 2013; Hayama et al. 2003; Kojima et al. 2002; Hill and Li 2016). In contrast to *A. thaliana*, rice *Hd1* represses *Hd3A* under LD condition whereas induces *Hd3A* function during SD period. Nevertheless, the florigen *Hd3A* is expressed in the leaf and then its protein is transported to the SAM where it interacts with OsFD (orthologoue of AtFD) to promote function of floral meristem identity gene *OsMADS15* (an orthologoue of AtAP1) similar to *A. thaliana* (Taoka et al. 2011). The interaction between Hd3a and OsFD in SAM is mediated by 14-3-3 proteins leading to formation of a florigen activation complex (FAC). Rice TFL1 like proteins namely RICE CENTRORADIALIS (RCN) is considered to have negative role in rice flowering similar to *A. thaliana* and it forms a florigen repressive complex (FRC) by interacting with 14-3-3 and OsFD in SAM in order to antagonize aforementioned FAC, thereby a balance between FAC:FRC is vital for optimal reproductive development (Kaneko-Suzuki et al. 2018).

Grain number, plant height and heading date 7 (*GHD7)* carrying a CCT domain act as a key regulatory gene in rice that recognizes day-length and it encodes a protein needed for grain number, plant height and heading date (Xue et al. 2008). *GHD7* downregulates expression of the B-type response regulator *EARLY HEADING DATE 1* (*EHD1*) in LD, but does not affect its expression under SD (Hill and Li 2016). *EHD1* is a blue light responsive promoter of flowering in rice that is also mediated by circadian clock (Doi et al. 2004; Andres and Coupland 2012). Expression of *EHD1* is upregulated by *OsGI* through *OsMADS51* under SD (Kim et al. 2007). Homologues of *GHD7*, *EHD1* and *OsMADS51* are absent in *A. thaliana* (Doi et al. 2004; Yang et al. 2012; Hill and Li 2016). *EHD1* promotes activity of *Hd3A* and *RFT1* in initiating flowering under SD and LD respectively (Hill and Li 2016).

1.3.2 Wheat (Triticum aestivum) and barley (Hordeum vulgare)

Temperate cereals such as wheat and barley are LD plants and responsive to vernalization for promotion of flowering, similar to *A. thaliana*. Adaptation of wheat and barley to various environmental conditions have been facilitated through allelic diversity in genes that mediate response towards photoperiod, e.g., *PPD* gene and growth practices, e.g., *VRN* genes (Distelfeld et al. 2009). Temperate cereals are divided into photoperiod sensitive or insensitive groups based on variation in *PPD* genes while winter and spring growth habit is determined by difference in *VRN* genes. *PPD1* is an orthologue of *A. thaliana* circadian clock gene *AtPRR7* (Turner et al. 2005; Hill and Li 2016). On the other hand, three *VRN* genes involved in the regulation of flowering are *VERNALIZATION 1* (*VRN1*) (a homologue of *A. thaliana AP1*/*FU*L) encoding a MADS box TF, *VRN3*/*FT1* (a homologue of *A. thaliana FT*), and *VRN2* which is a zinc finger protein with a CCT domain having no clear homologoue in *A. thaliana* (Yan et al. 2003; Yan et al. 2004; Yan et al. 2006). Regulation of *VRN3*/*FT1* is a point where photoperiod and vernalization pathways traverse with each other (Yan et al. 2006; Trevaskis et al. 2007; Hill and Li 2016).

Nine orthologues of *CO* have been identified in barley with *HvCO1* and *HvCO2* having highest similarity to *AtCO* (Griffiths et al. 2003). In case of wheat, two *CO* like genes namely *TaHd1* and *WCO1* have been identified (Shimada et al. 2009). Sequences having homology with *AtGI* have already been identified in both barley and wheat (Dunford et al. 2005; Zhao et al. 2005). It has been proposed that a photoperiod mediated GI-CO-FT module is likely to occur in wheat TaGI-TaHd1/WCO1-VRN3 (Shimada et al. 2009). *HvGI* has a similar expression pattern to that

of *AtGI* under LD and is assumed to act in a similar GI-CO-FT module based on comparative studies using *Brachypodium distachyon*, but specific role of *HvGI* with regard to mediating *HvCO1*/*HvCO2* expression have not been elucidated yet (Dunford et al. 2005; Higgins et al. 2010; Alqudah et al. 2014). Nevertheless, *HvCO1* is known to induce expression of *VRN3*/*FT1* under both LD and SD (Campoli et al. 2012). *PPD1* acts upstream of *CO* to promote flowering in case of both barley and wheat (Distelfeld et al. 2009).

Similar to *A. thaliana*, VRN3/FT1 act as a mobile protein in wheat and barley which moves from leaf to SAM where it interacts with FDL2 (*A. thaliana* FD like protein) resulting in upregulation of *VRN1*/*AP1*. Vernalization promotes *VRN1*/*AP1* expression which then repress *VRN2*, thereby flowering is initiated (Distelfeld et al. 2009; Hill and Li 2016). Thus, the genetic process driving vernalization response in wheat and barley is distinct from *A. thaliana* where FLC like proteins yet to show similar functions (Ruelens et al. 2013; Hill and Li 2016).

In higher latitude regions, spring barley cultivars were selected based on a recessive mutation in the CCT domain of *ppd1*, which resulted in late flowering and maturity under LD conditions. These traits assisted in adaptation of barley cultivars under long summer growing seasons of Central and Northern Europe (Turner et al. 2005; Campoli and Von Korff 2014). In contrast, day-length neutrality has been obtained in wheat by dominant mutations in *PPD1* homeologs, resulting in early flowering under both SD and LD conditions. This trait has special adaptive importance for yield benefits of wheat grown in Central European countries (Worland et al. 1998) and Australia (Richards et al. 2014).

Barley flowering time loci *early maturity 8* (*eam8*) and *eam10* have been recently identified as orthologues of circadian clock genes *ELF3* and *LUX,* respectively (Campoli et al. 2013; Faure et al. 2012; Zakhrabekova et al. 2012)**.** Lines carrying non-functional eam8 or eam10 protein show elevated *PPD1* expression and subsequent induction of *FT1* expression and acceleration of flowering (Campoli et al. 2013; Faure et al. 2012). In addition, mutation in a *LUX*-like gene also result in elevated expression of *PPD1* and *FT1* and early flowering under non-inductive SD conditions (Mizuno et al. 2012). Therefore, it could be stated that *eam* loci may have assisted in adaptation of barley and wheat in certain environments such as high-altitude short-growing seasons of Europe and it led to the generation of new variety (Campoli and Von Korff 2014).

1.3.3 Maize (Zea mays)

Maize is a facultative SD plant and many of the *A. thaliana* orthologues have been found to regulate flowering in this plant. Signals from photoperiod, circadian clock, age, autonomous and gibberellin-mediated pathway are integrated by different genes (Dong et al. 2012) to control flowering in maize. Similar to *A. thaliana*, FT homologue in maize named as Zea mays CENTRORADIALIS8 (ZCN8) moves from leaf to shoot apex to interact with FD homologue called Delayed Flowering1/DLF1 to induce function of the maize floral identity gene *ZMM4* which is orthologue of *AtFUL1* (Meng et al. 2011; Wigge et al. 2005). Apart from *ZCN8*, a second florigen *ZCN7* have also been proposed in maize (Mascheretti et al. 2015; Hill and Li 2016). Other genes such as meristem identity genes *ZAP1a* (*Arabidopsis* orthologue *AP1*) and *ZAP1b* are expressed in the apex and play positive role in floral development (Heuer et al. 2001; Mena et al. 1995). Likewise, *ZFL1* (*Arabidopsis* orthologue *LFY*) and *ZFL2* promote flowering at the apex (Bomblies et al. 2003).

The *CONZ1/ZmCO1* is orthologous to *A. thaliana* flowering gene *CO* and it plays an important role in positively regulating flowering via incorporation of photoperiodic and clock signals. Unlike *A. thaliana*, interaction between CONZ1 and ZCN8 is still elusive (Dong et al. 2012; Miller et al. 2008; Hill and Li 2016). *ZmCCT1* is an orthologue of rice *Ghd7* which integrates signals from photoperiod and circadian clock to negatively regulate function of *ZCN8* (Hung et al. 2012). *INDETERMINATE1* (*ID1*) is a cereal specific gene expressed in the leaf and it assists in the movement of *ZCN8* from leaf to shoot apical meristem via phloem (Coneva et al. 2007)

1.3.4 Sorghum (Sorghum bicolor)

Sorghum is a SD plant. Grain sorghums have reduced sensitivity to photoperiod and therefore breeders chose genotypes with such traits for growing in the temperate region in order to avoid drought and unfavorable temperature and produce more grains. On the contrary, energy sorghums are highly photoperiod sensitive and they undergo long vegetative growth before being able to flower (Olson et al. 2012; Rooney et al. 2007).

Similar to *A. thaliana*, the regulation of flowering time via GI-CO-FT is conserved in sorghum and corresponding homologs have been identified. *SbCN8* and *SbCN12* are homologues of *A. thaliana FT*/maize *ZCN8* and *ZCN12* respectively and act as florigen in sorghum (Murphy et al. 2011; Yang et al. 2014b). However, unlike *A. thaliana*, six maturity loci namely *Ma1-Ma6* were

identified in sorghum and dominance at these loci slows down flower initiation (Morgan and Finlayson 2000; (Rooney and Aydin 1999). Sorghum also possesses *EHD1*, *GHD7* and *PRR37* gene similar to rice, barley and wheat which play vital role to regulate flowering (Yang et al. 2014a).

1.3.5 Soybean (Glycine max)

Soybean is a major legume crop and has become an important model system for investigation of mechanisms controlling flowering in warm-season SD legumes (Weller and Ortega 2015). Unlike *A. thaliana*, flowering time and maturity is under the control of nine major genes, termed *EARLY MATURITY, E1 - E9*. The majority of these genes delay flowering under non-inductive LD condition, and derived recessive genotypes are early-flowering (Kong et al. 2014; Watanabe et al. 2012). In comparison to *A. thaliana*, soybean carries 10 *FT* genes that fall into three major clades; *FTa*, *FTb* and *FTc* (Laurie et al. 2011). The most prominent of these genes are *FT2a* (an *FTa*-type gene) and *FT5a* (an *FTc*-type gene). Under SD, *FT2a* and *FT5a* gene expression is induced in the leaves resulting in initiation of floral formation (Weller and Ortega 2015). The *FT4* gene (an *FTb*-type gene) which in contrast to most other known *FT* genes, acts as a repressor of flowering (Zhai et al. 2014), a distinctive role whose conservation across other legumes is not yet clear. Similar to *A. thaliana*, soybean orthologues of *PHYA* and *GI* play important role in regulating photoperiodic flowering, although their mechanisms of action are not yet clear (Watanabe et al. 2009; Watanabe et al. 2012). Four homologues of *A. thaliana CO* have been discovered in soybean (Wong et al. 2014) and are able to promote flowering in transgenic *A. thaliana* (Wu et al. 2014). However, role of *CO* is yet to be fully understood in soybean and it is assumed that they are not the major regulator of flowering. This role may instead be played by *E1* a legume-specific gene that is controlled by light and circadian clock whose function is dependent on the *PHYA* gene *E4* (Xia et al. 2012). E1 protein consists of a putative bipartite nuclear localization signal and a part indirectly associated with B3 domain. *E1* negatively modulates function of *FT2a* and *FT5a* (Thakare et al. 2011; Xia et al. 2012) whereas it promotes function of *FTb4* (Zhai et al. 2014). Thus, *E1* acts as a key repressor of the flowering pathway in soybean.

1.4 Pea (*Pisum sativum*)

1.4.1 An ideal legume species

Pea is an important crop species that similar to other annual legumes is ecologically beneficial due to its ability to fix atmospheric nitrogen as well as potentiality to serve as a break crop. Pea seeds are good source of specially starch and proteins and contain minor proportion of soluble sugars, fiber, minerals and vitamins (Bastianelli et al. 1998). Due to its short generation time, ability to reproduce via self or cross-pollination, suitability of grafting and availability of diverse lines, pea serves as a model plant species for molecular genetics studies related to flowering time. Pea is cultivated worldwide with Canada, China, France, India, Russia and Australia currently being the top producing countries (Figure 1.5).

For plant biologists, genetic diversity as naturally occurring variation and induced mutagenesis serve as a source of material for elucidating molecular mechanism regulating flowering time in pea. Due to the advent of various molecular biology techniques such as molecular marker development during the past 10-15 years resulted in significant upsurge in identification of flowering time genes in pea. Comparative mapping studies with other model species also facilitated research in this direction. Barrel medic (*Medicago truncatula*) is taxonomically the closest plant species to pea and genomic sequence data is available for this plant species. Therefore, *M. truncatula* serves as a good resource in functional genomics

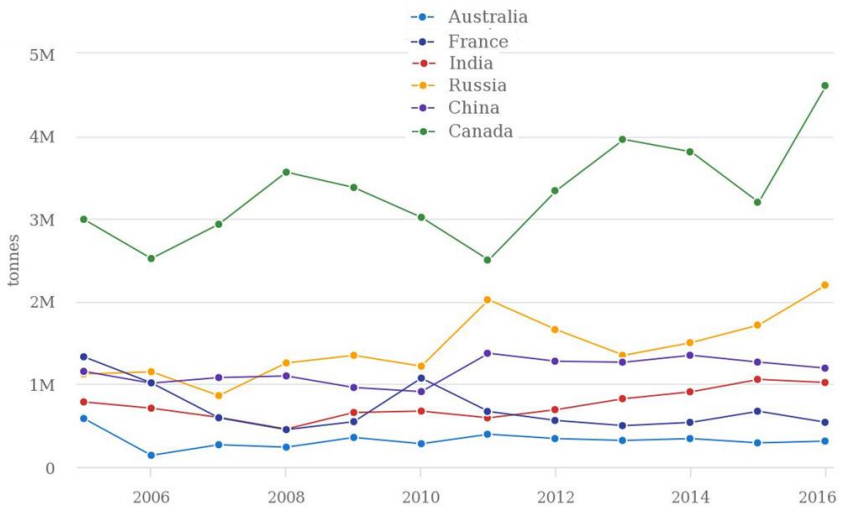


Figure 1.5. Production of dry pea in top five productive countries from 2005-2016. Source: http://www.fao.org/faostat/en/#compare

studies and candidate gene identification in pea with relevance to flowering time. Even though genomic sequence of pea is not yet available, transcriptomic database have already been developed by relevant scientific community (Susete et al. 2015).

Pea (*Pisum sativum*) is a LD plant and similar to *A. thaliana* is responsive to vernalization. Unlike *A. thaliana* and similar to soybean, pea contains multiple *FT* genes in the *FTa*, *FTb* and *FTc* subclades (Hecht et al. 2011). Among the *FT* genes, *FTa1* and *FTb2* are suggested to act as mobile elements that transport flowering signal from the leaves to the shoot apical meristem (SAM) under LD condition (Weller and Ortega 2015). Besides, *CO* has no known role in regulating flowering in pea.

Isolation and characterization of induced mutants has been a key research strategy to identify genes responsible for flowering in pea. In this way, the role of several pea homologues of known *A. thaliana* flowering genes have been unveiled. Based on their phenotype of floral initiation, these been grouped either as early flowering or late flowering mutants.

1.4.2 Early flowering mutants

The early flowering mutants that have been functionally characterized successfully in pea are *det* (Foucher et al. 2003), *lf* (Foucher et al. 2003), *dne* (Liew et al. 2009), *sn* (Liew et al. 2014), *phyB* (Weller et al. 2001) and *hr* (Weller et al. 2012a). Three such genes are described below:

1.4.2.1 late flowering (lf) and determinate (det)

In *A. thaliana*, *TERMINAL FLOWER1 (TFL1)* mediates inflorescence meristem identity and floral initiation by repressing flowering (Ratcliffe et al. 1998). Thus, *TFL1* controls two distinct functions by regulating the duration of both vegetative and reproductive phases. The *tfl1* mutant in *A. thaliana* has determinate inflorescence with a terminal flower and possess early flowering phenotype compared to the wild type (Bradley et al. 1997).

LATE FLOWERING (LF) was the first key flowering gene to be discovered in pea (White 1917). *LF* is a repressor of flowering and it governs the flowering node in a given pea accession. *LF* slows down transition from vegetative to inflorescence meristem which results in delayed floral initiation (Figure 1.6). Four natural and induced types of alleles has been identified namely *Lf-d, Lf, lf* (Figure 1.7) and *lf-a* for which minimum node of flowering is 15, 11, 8 and 5 respectively.

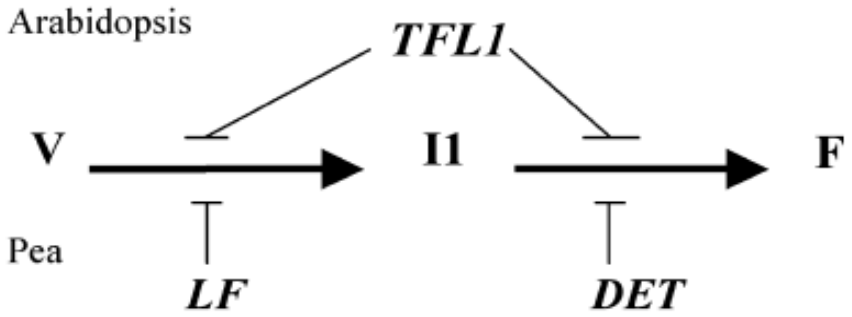


Figure 1.6. Role of *TFL1* genes to regulate floral initiation in *A. thaliana* and pea. Source: (Foucher et al. 2003). F- flower; I_1- inflorescence meristem; V-vegetative phase.

Phylogenetic analysis has revealed the presence of three homologues of *TFL1* in pea namely *TFL1a*, *TFL1b* and *TFL1c*. Through gene mapping, allele sequencing and expression studies one of these homologues *TFL1c* was identified as *LF* (Foucher et al. 2003).

DET regulates floral phase in pea by inhibiting the transition from inflorescence meristem to flower (Figure. 1.6). *DET* has been identified as *TFL1c*, which is one of the three homologs of *A. thaliana TFL1* (Foucher et al. 2003). The *det* mutant shows determinate growth of main apex without influencing flowering time. Growth of the shoot ceases in *det* mutants after generation of few reproductive nodes. In this case, primary inflorescence is transformed early to secondary inflorescence, which terminates into a stub after giving rise to one or two flowers (Figure 1.7) (Benlloch et al. 2007). Thereby, the main shoot is shifted into an organ with restricted growth instead of unlimited growth of the wild-type plants. Expression of *DET* occurs only in the shoot apex once floral transition is initiated.

1.4.2.2 sterile node (sn)

STERILE NODE (SN) is one of the earliest studied locus that was used to understand photoperiod response in pea (Barber 1959; Murfet 1971). *SN* mediates photoperiodic sensitivity and plants carrying dominant *SN* locus generates higher number of vegetative nodes compared to the wild type and exhibit late flowering phenotype (Barber 1959; Murfet 1971). Recessive mutation at the *SN* locus results in day-length insensitivity and reduced vernalization response as well as early flowering. Besides, three different induced mutants carrying alleles of *sn* mutation have also been discovered. These are: *sn-2* generated via

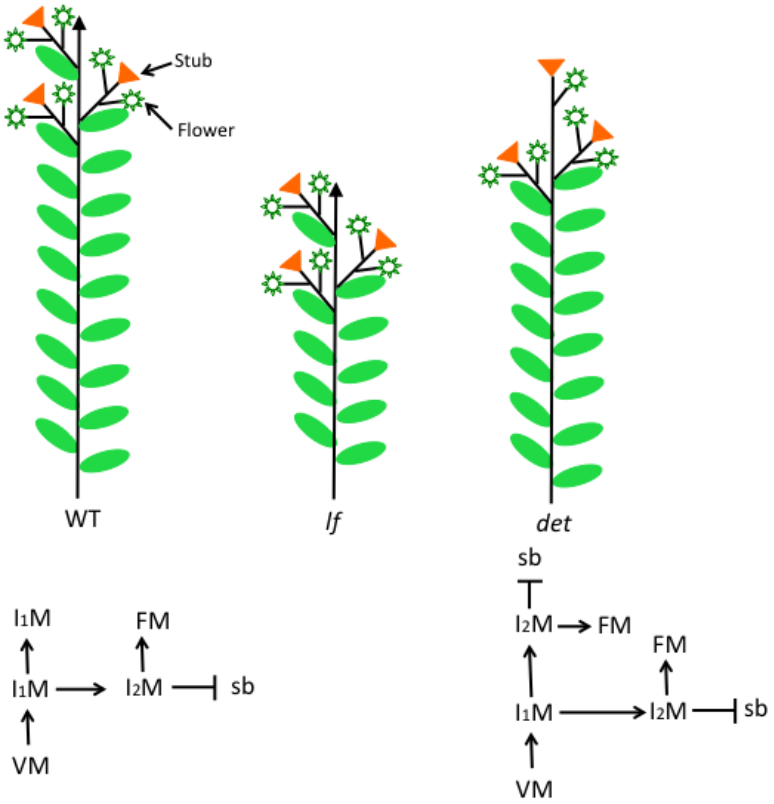


Figure 1.7. Schematic diagram representing flowering phenotype of WT, *lf* and *det* mutants in pea. VM= vegetative meristem, I_1M= primary inflorescence meristem, I_2M= secondary inflorescence meristem, FM= floral meristem, sb=stub. Figure adapted from Sussmilch F., 2014.

radiation mutagenesis in cv. Borek background (Arumingtyas and Murfet 1994) and sn-3 and sn-4 developed through EMS mutagenesis in NGB5839, cv. Torsdag background (Hecht et al. 2007; Liew et al. 2014).

SN locus has been functionally characterized and identified as the orthologue of *A. thaliana* circadian clock gene *LUX* (Liew et al. 2014). Together with three other photoperiod response genes *DAY NEUTRAL/DNE* (pea homolog of *ELF4*) (Liew et al. 2009), *HIGH RESPONSE TO PHOTOPERIOD/ HR* (pea homolog of *ELF3a*) (Weller et al. 2012a) and *PHOTOPERIOD RESPONSE*/*PPD* (pea homolog of *ELF3b*) (Rubenach et al. 2017), *SN* forms the evening complex of circadian clock. *HR*/*PPD*/*DNE* are considered to have redundant role for initiating flowering and genetic interaction studies suggested that *HR* and *DNE* likely to promote *SN* activity in order to repress flowering (Rubenach et al. 2017; Liew et al. 2014). Mutation in these four

genes likely to have played important role in adaptation of pea in shorter growing seasons (Liew et al. 2009; Weller et al. 2012a; Liew et al. 2014; Lu et al. 2017; Rubenach et al. 2017).

1.4.2.3 lv (phyB)

Six different mutant alleles of the *LV* gene were generated through different processes in different genetic backgrounds. Among them, *lv-1* and *lv-2* mutants were generated in the background of a pea cultivar named Sparkle via application of nitroso ethyl urea and gamma radiation (Reid JB 1988). Besides, *lv-3* was created through fast neutron from cv. Paloma and *lv-4* mutant originated spontaneously from Hobart line 80/L80 (Weller et al. 1992). While *lv-1* showed presence of PHYB apoprotein, opposite result was obtained for the *lv-2*, *lv-3* and *lv-4* mutant alleles (Weller et al. 1995). The *lv-1* and *lv-2* mutant exerted strong response for traits such as leaf area and plant height when grown under continuous red light. Besides, *lv-1* and *lv-4* showed response for shade-avoidance symptoms (leaf area, plant height) under low R:FR (ratio = 0.66) compared to high R:FR conditions (ratio = 5.44). Two additional mutant alleles of *LV* gene namely *lv-5* and *lv-6* were obtained through EMS mutagenesis of cv. Torsdag (Weller et al. 2001). PHYB apoprotein could not be detected in *lv-5* whereas it was present in *lv-6* similar to *lv-1*. Phylogenetic analysis confirmed that *LV* is the pea homologue of *PHYB* and multiple, independent functional mutation for this gene was identified by cDNA sequencing of *PHYB* in *lv-1*, *lv-5* and *lv-6* alleles (Weller et al. 2001). *PHYB* was found to inhibit flowering under both long and short day conditions and this effect was not transmissible from leaf to shoot apex.

1.4.3 Late flowering mutants

Several late flowering mutants namely *fun1* (Weller et al. 1997a; Weller et al. 2004), *late1* (Hecht et al. 2007), *gigas* (Hecht et al. 2011), *veg2* (Sussmilch et al. 2015) and *late2* (Ridge et al. 2016) have so far been functionally characterized in pea.

1.4.3.1 fun1 (phyA)

The *fun1* mutant were generated in the background of cv. Torsdag background via EMS mutagenesis which exhibited photoperiod insensitive flowering initiation (Weller et al. 1997a). Immunoblotting experiment revealed that *fun1* mutants lack Phytochrome A/PHYA apoprotein. Another *phyA* mutant namely *AFO5*/*phyA-3D* possessing dominant inheritance with early flowering phenotype under SD was also identified by another mutant screening

(Weller et al. 2004). Genetic mapping and candidate gene sequence analysis confirmed that *fun1, fun2* and *AF05* are allelic variants of *PHYA* gene (Weller et al. 2004). The *fun1* mutants showed no difference in flowering under low intensity (8 µmol m^{-2} s^{-1}) and high intensity white fluorescent light conditions (150 µmol m^{-2} s^{-1}) devoid of far-red light. These results suggested that *PHYA* in pea is involved in regulating photoperiod response under both high and low R : FR ratios. Likewise, grafting and genetic interaction studies involving *fun1* and *dne* (deficient in floral inhibitor) showed that promotion of flowering regulated by *PHYA* is a result of reduced synthesis or transmission of a floral inhibitor. Moreover, genetic interaction studies suggested that *PHYA a*nd *PHYB* function redundantly to control seedling de-etiolation under continuous red light and *PHYB* acts after *PHYA* for regulating flower initiation (Weller et al. 2001).

1.4.3.2 late bloomer1 (late1)

Six recessive mutant alleles have been discovered for *LATE1* locus namely *late1-1* to *late1-6*, all of which have traits such as significantly delayed flowering compared to the wild type, increased basal branching and prolonged reproductive phase (Hecht et al. 2007). Functional genomics studies of the *LATE1* locus revealed that it is an orthologue of *A. thaliana GI* gene. This photomorphogenic effect of *LATE1* is dependent on PHYB. Likewise, flower-inducing function of *LATE1* also rely on active PHYB. In contrast, *LATE1* act *PHYA* act additively in order to facilitate reproductive development (Hecht et al. 2007).

In addition, *late1* mutants show disruption in the rhythmic expression of circadian clock related genes such as *LHY* and *DNE*. Therefore, *LATE1* is crucial for proper functioning of key circadian clock genes (Liew et al. 2009). *LATE1* has also been proposed to promote function of key *FT* genes in pea namely *FTa1* and *FTb2* in the leaf (Hecht et al. 2011). Moreover, grafting experiments elucidated that *LATE1* is involved in the synthesis of a mobile flowering stimulus in pea.

1.4.3.3 gigas (fta1)

Three recessive mutant alleles for the *GIGAS* locus have been defined. Among them, *gigas-1* was selected upon gamma radiation of the late flowering, long day variety Virtus (Beveridge and Murfet 1996). The second allele *gigas-2* was the result of fast-neutron mutagenesis of cultivar Porta whereas the third allele *gigas-3* was generated via EMS mutagenesis of line

NGB5839 (Murfet 1992; Taylor and Murfet 1994; Hecht et al. 2011). Under SD, these mutants show delayed flowering compared to WT (Weller 2007; Beveridge and Murfet 1996; Reid et al. 1996). In contrast, these mutants showed marked difference in various phenotypes including flowering time under LD (Weller 2007). Mutant allele exhibiting stronger phenotype such as *gigas-2* do not flower at all under LD (Hecht et al. 2011). Grafting experiment using *gigas* mutant scion onto WT stock showed that normal flowering can be rescued in this mutant which gave hints that *GIGAS* gene is involved in generation of graft-transmissible flowering promoting mobile signal and FT like genes are likely candidates for this mutation (Beveridge and Murfet 1996; Weller 2007; Hecht et al. 2011). Indeed, genetic mapping and candidate gene analysis revealed that pea *FTa1* gene corresponded *GIGAS* locus (Hecht et al. 2011). Besides, gene expression analysis unravelled that *GIGAS*/*PsFTa1* promote expression of pea inflorescence identity genes *PIM*, *SEP1* and *UNI* under LD (Hecht et al. 2011). Likewise, this gene was also found to be essential for full expression of *FTa2* in leaf and apex and *FTc* in apex. These results established the potentiality of mutual regulation between *FT* genes which is key for floral induction in pea. In addition, it was suggested via genetic interaction studies that *GIGAS* positively regulate flowering by repressing function of *LF/PsTFL1c* which is a negative regulator of flowering (Hecht et al. 2011).

1.4.3.4 vegetative2 (veg2)

Two recessive mutant alleles namely *veg2-1* and *veg2-2* were generated by fast neutron mutagenesis (Murfet 1992, Murfet and Reid 1993). Among these, *veg2-1* does not form any flower during its entire development growth whereas *veg2-2* flower significantly later than the wild-type. Both these mutants also have characteristic features like formation of greater numbers of aerial branching, which occupy aerial nodes. Thus, vegetative to primary inflorescence (I_1) formation step is highly affected in these mutants. *VEG2* has been identified as the *A. thaliana* orthologue of *FD* (Sussmilch et al. 2015).

Grafting studies revealed that *VEG2* regulates flowering by acting locally at shoot apex (Sussmilch et al. 2015). Besides, it plays decisive role is defining and establishing secondary inflorescence (I_2) identity. Ubiquitous expression pattern of *VEG2* at the apex has been observed and it was found to be expressed in vegetative shoot apical meristem, axillary meristem, I_1 meristem, I_2 meristem, vasculature, tips of leaf primordial and early stages of floral meristems. Therefore, *VEG2* is a key regulator of floral meristem identity. VEG2 interacts

physically with FT proteins and therefore it is assumed that VEG2-FT complex is essential for floral promotion. Likewise, various flowering and meristem identity genes in pea such as *FTa1*, *FTc*, *DET*, *LF, PIM*, *VEG1* are misregulated in *veg2* mutants that indicate dependence of these genes on *VEG2* for proper regulation and thereby mediating flowering (Sussmilch et al. 2015).

1.4.3.5 late bloomer2 (late2)

During an EMS mutant population screening in the background of pea line NGB5839 grown under LD conditions, the *late2-1D* mutant was identified which showed features similar to photoperiod response mutants such as delayed flowering, prolonged reproductive phase and formation of basal branching (Hecht et al. 2007; Ridge et al. 2016). Dominant inheritance of these phenotypes were observed among the F_2 population generated via cross between progenitor NGB5839 and *late2-1D* (Ridge et al. 2016). Genetic interaction studies involving other photoperiod response mutants such as *late1-2* and *phyA-1* showed delayed flowering and higher numbers of reproductive nodes in the double mutants which was in line with the fact that all the three underlying genes act in the same pathway. Grafting experiments confirmed that *LATE2* gene is involved in the process of production of graft transmissible signal.

Genetic mapping, phylogenetic analysis and positional candidate gene based approach unveiled molecular identity of *LATE2* gene as legume *DOF* transcription factor encoding gene *CDFc1* (Ridge et al. 2016). Further research involving yeast two hybrid gave strong evidence that mutation in the *CDFc1* gene of *late2-1D* affect its physical interaction with FKF1 which is needed for the function of this complex with relevance to regulating flowering. Likewise, transformation of *A. thaliana* with mutant form of *PsCDFc1* from *late2-1* exhibited late flowering phenotype in various transformed lines, thus it was established that mutation in *PsCDFc1* gene in *late2-1* is responsible for the delayed flowering phenotype. Gene expression experiment unravelled that a FKF1/GI/CDF module mediates expression of *FT* genes in pea and they also regulate photoperiodic response. But, orthologues of *CO* are not targets of this module which led to the proposal of hypothesis that a different underlying process forms bridge between FKF1/GI/CDF and *FT*s (Ridge et al. 2016).

1.4.3.6 Unknown late bloomer3 (late3) and late bloomer4 (late4)

EMS mutant screening led to the discovery of two extremely late flowering mutants namely *late3* and *late4* at the School of Natural Sciences, University of Tasmania (Weller 2007). These were generated via EMS mutagenesis in the background of wild-type pea line NGB5839. The *LATE3* locus has three recessive mutant alleles known as *late3-1, late3-2* and *late3-3* whereas two recessive mutant alleles have been identified for the *LATE4* locus called *late4-1* and *late4-2*.

Due to their clear phenotyping difference in flowering compared to the wild type, these mutants are of great interest for scientists for functional characterization studies and thereby reveal the identity and function of the underlying gene with relevance to flowering time. More information on phenotypes of these mutants have been provided in chapter 3.

1.5 Aims for this study

The overall goal of the current study was to functionally characterize two novel late flowering pea mutants namely *late3* and *late4* for which the genes have not yet been identified. Therefore, various investigations were undertaken in this project in order to reveal the identity of the *LATE3* and *LATE4* genes and their role in relevance to flowering in pea. As part of these aims, chapter 3 and 4 describe phenotypic characterization of *late3* and *late4* mutants and mapping of *LATE3* and *LATE4* leading to identification of the chromosomal position in pea where these genes are located. This was followed by candidate gene analysis in chapter 5 that revealed identity of the genes function of which have been impared in the *late3* and *late4* mutants. On the other hand, chapter 6 largely explains the role of *LATE3* and *LATE4* genes in regulating flowering in pea through genetic interaction with known early flowering pea mutants and regulatory interaction with important pea flowering genes. Finally, chapter 7 elucidates role of *LATE3* and *LATE4* genes in various biological processes by studying their responsiveness to different abiotic factors.

Chapter 2: General materials and methods

Various general materials and methods that were used in current study have been mentioned in this chapter. Detailed information on materials and methods for specific experiments have been provided in the materials and methods section of respective chapters.

2.1 Plant materials

Various pea genotypes used in this thesis are given in table 2.1. Specific purpose of using these genotypes are mentioned in materials and method section of research chapters.

Table 2.1. Details on different pea genotypes used in the current study

Genotype	Background	Description	Reference
NGB5839	Torsdag	EMS mutant, dwarf due to mutation in GA 3β-hydroxylase gene *LE*	Lester et al. 1997
Terese	-	Cultivated pea line, leaves transformed into tendrills due to mutation in leaf morphogenesis *AF* gene	Lorindon et al., 2005
late3-1	NGB5839	EMS mutant, late flowering under LD	Weller 2007
late3-2			
late3-3			
late4-1			
late4-2			
lf-22		EMS mutant, early flowering under LD	Weller et al., 2007
sn-4			Hecht et al., 2007
det-2		Spontaneous mutant backcrossed 3 times to NGB5839, early flowering under LD	Foucher et al., 2003

2.2 Growth conditions

For all the experiments involving pea, the pots that were used had a size of 14 cm. Before sowing, all the pots were sterilized with 90% ethanol and filled initially with 1:1 mixture of dolerite chips and vermiculite which was topped with 3 cm of sterile native nursery grade potting mix containing controlled release fertilizer (CRF) (Horticultural and Landscape

Suppliers, Brighton, Tasmania, Australia). The seeds were coated with fungicide powder Thiram prior to sowing. Usually, seeds were placed at a depth of about one knuckle of the index finger. Pots were watered daily until seed germination and it was followed by thrice/ week until flowering and again daily until harvesting. Plants were given 0.1% Peters professional nutrients once a week.

Plants were grown in three different facilities within the School of Natural Sciences, University of Tasmania based on the goal of the experiment as well as availability of space. Name and growth conditions of these facilities are as follows:

i. Glass house – 18 hour photoperiod, 20° C, natural day light in the morning and afternoon is extended by using high pressure sodium 400 W in Pierlite (Padstow, NSW, Australia) HL400HPSCG pendant mount fittings with acrylic diffusing covers delivering nearly 100 µmol m^{-2} s^{-1} at the pot surface.
ii. Top phytotron – 16 hour photoperiod, 20° C, natural day light in the morning and afternoon is supplemented with the same artificial light conditions mentioned above.
iii. Controlled growth chamber - 16 hour photoperiod, 20° C, white light devoid of far-red spectrum provided by cool white fluorescence tubes (L40 W/20 S cool white, Osram, Germany) at an intensity of 120-140 µmol m^{-2} s^{-1}.

2.3 Recording phenotypic data

2.3.1 Node of flower initiation, node of flower development, total node and reproductive node

The node in which first floral bud appeared along the main stem is called node of flower intiation (NFI) where first scale leaf carrying node is considered as node 1. Similarly, the node at which first open flower appeared is considered as the node of flower development (NFD).

Number of total nodes (TN) were counted from six plants of NGB5839, *late3-1, late3-2, late3-3, late4-1* and *late4-2* genotypes at the time of harvest (180 days after sowing)

The number of reproductive nodes (RN) were calcluated for the wild-type during the time of harvest according to following formula: RN = TN – NFI + 1). For all the mutants, initial reproductive nodes (I-RN) were counted for six plants at the time of harvest (180 days after

sowing). If inflorescence reversion was observed, then the number of vegetative nodes after the I-RN were counted and these nodes were termed as inflorescence reversed vegetative nodes (IR-VN). In the case where new reproductive nodes were observed after IR-VN, then these nodes were counted and they were given the term later reproductive nodes (L-RN).

2.3.2 Total number of flowers

Total number of flowers were counted which were produced both in main stem and lateral branches were counted from six plants of NGB5839, *late3-1, late3-2, late3-3, late4-1* and *late4-2* genotypes at the time of harvest (180 days after sowing).

2.3.3 Pod traits

Total number of seed carrying pods produced both in main stem and lateral branches were counted from six plants of NGB5839, *late3-1, late3-2, late3-3, late4-1* and *late4-2* genotypes at the time of harvest (180 days after sowing). These pods were also sorted according to the number of 5, 4, 3, 2 and 1 seed carrying pods from each of the genotypes.

Percentage of seed carrying pods was calculated for each of these genotypes by following formula: (Total numbers of seed carrying pods X 100)/Total number of flowers.

2.3.4 Total number of seed and dry weight of seeds

Total number of seeds were counted which were produced in various types of pods from six plants of NGB5839, *late3-1, late3-2, late3-3, late4-1* and *late4-2* genotypes at the time of harvest (180 days after sowing).

Dry weight of seeds were measured in mg for NGB5839, *late3-1, late3-2, late4-1* and *late4-2* genotypes by measuring the weight of 10 seeds from each genotype in three independent measurements.

2.3.5 Total branching

Any lateral outgrowth at a certain vegetative node >5 cm was considered as a lateral branch. Total branching was calculated from measurements of six plants of NGB5839, *late3-1, late3-2, late3-3, late4-1* and *late4-2* genotypes as a ratio between total length of all lateral branches and total length of stem at the time of harvest (180 days after sowing).

2.3.6 Leaf formation rate

Leaf formation rate was measured from 6-10 plants of NGB5839, *late3-1, late3-2, late3-3, late4-1* and *late4-2* genotypes starting from day 21 until first flower appeared in each genotype. Each fully expanded leaf was counted as 1 and unopen leaves were counted as 0.1-0.9 based on the status of their openness at the time of counting.

2.3.7 Leaf size

Leaf area was measured as product of width and length from 9-12 plants of NGB5839, *late3-2* and *late4-1* genotypes for single leaflet of leaf at 8th node when leaves still had two leaflets.

2.3.8 Petiole and proximal rachis length

Petiole and proximal rachis length were measured in cm for the 10th leaf from 6-10 plants of NGB5839, *late3-1, late3-2, late3-3, late4-1* and *late4-2* genotypes.

2.3.9 Stem diameter

Stem diameter was measured in mm from 6-10 plants of NGB5839, *late3-1, late3-2, late3-3, late4-1* and *late4-2* genotypes for the internode between node 10 and node 9 using a digital caliper (https://www.wpiinc.com/products/laboratory-supplies/501601-digital-caliper/). Since stem did not have completely round shape, therefore three independent measurements were taken at top, middle and bottom of 9th internode placing the caliper on the thickest part of the stem at 49 days after sowing.

2.3.10 Stem length

Stem length was measured in mm from six plants of NGB5839, *late3-1, late3-2, late3-3, late4-1* and *late4-2* genotypes for the length between node 1 and apex at the time of harvest (180 days after sowing).

2.3.11 Internode length

Internode length was measured in mm from six plants of NGB5839, *late3-1, late3-2, late3-3, late4-1* and *late4-2* genotypes for each internode at the time of harvest (180 days after sowing). Internode 1 represents the length between first and second scale leaves.

2.3.12 Primary root length and root dry weight

Length of primary root was measured in mm for 8-12 plants of NGB5839, *late3-2* and *late4-1* genotypes as the length between cotyledon and primary root tip 12 days after sowing.

Root dry weight was measured in mg for 8-12 plants of NGB5839, *late3-2* and *late4-1* genotypes by washing and harvesting the entire fresh root biomass below cotyledon after 12 days of sowing which was then dried at 70° C for 4 days.

2.3.13 Shoot fresh and dry weight

Fresh weight of shoot was calculated from 9-12 plants of NGB5839, *late3-2* and *late4-1* by harvesting the samples after 27 days of growth and immediately weighing them.

Dry weight of shoot was calculated by harvesting the samples after 27 days of growth then drying at 70° C for 4 days.

2.4 Primer design

Various primers that were used for research involving pea were generated by using the software Geneious 8.0.4 (Kearse et al. 2012). As part of that structure of orthologous *Medicago truncatula* transcript was collected from *M. truncatula* Mt4.0 v1 genome database (https://phytozome.jgi.doe.gov/pz/portal.html#) and aligned with pea transcript sequences obtained from pea transcriptome database (http://bios.dijon.inra.fr/FATAL/cgi/pscam.cgi) in order to annotate pea transcripts. All the primers that were used in this study for different purposes have been mentioned in tables A1.1-A.4.

2.5 DNA extraction

In order to extract DNA, leaf tissues samples were placed into liquid nitrogen immediately after harvesting and stored at -70° C until processing. These tissue samples were ground using carbide beads in a mechanical homogeniser (Retsch MM30 or Qiagen Tissue Lyser II) for 2 minutes at 30 r/s. Then, 500 µl of extraction buffer (100mM Tris-HCl pH 8.0, 1.4 M NaCl, 20 mM EDTA, 2% w/v CTAB) was added to the powdered tissue and shaken in the mechanical homogenizer for 1 minute at 30 r/s. Afterwards, samples were incubated for 15 minutes 60° C. It was followed by addition of 500 µl of Chloroform : Isoamylic alcohol (24 : 1) and gentle mixing by hand. The samples were then centrifuged at 14000 rpm for 5 minutes and the upper aqueous phase was taken out into new tubes and extracted for the 2nd time. Precipitation of the DNA was carried out by adding 1 ml of precipitation buffer (50 ml Tris-HCl pH8.0, 10 mM EDTA, 1% w/v CTAB) and incubating for 10-15 minutes at room temperature after gentle mixing which was followed by centrifugation at 14000 rpm for 10 minutes. The precipitate

was resuspended in 300 µl of 1.5 M NaCl containing 1 µl of RNase A (25 mg/ml) and incubated for 1 hour at 50° C. DNA was then pelleted by adding 600 µl of 95% ethanol and incubating for 5-10 minutes at room temperature followed by centrifugation at 14000 rpm for 10 minutes. The pellet was washed with 200 µl of 70% ethanol and centrifuged for 5 minutes at 14000 rpm. Finally, the washed pellet was dried at 37° C for nearly 20 minutes and eluted with 50 µl of sterilized distilled water and kept at 4° C overnight.

Concentration of the extracted DNA was measured in Nanodrop 800 Spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA) and diluted to 50 ng/µl concentration for PCR.

2.6 RNA extraction and cDNA synthesis

As part of RNA extraction, frozen leaf tissue samples were ground using mortar and pestle whereas apex tissue samples were crushed using carbide beads in a mechanical homogeniser (Retsch MM30 or Qiagen Tissue Lyser II) for 2 minutes at 30 r/s. Extraction of total RNA was performed using Promega SV Total RNA Isolation System (Promega, Madison, WI) by following manufacturer's instruction. Quantification of the extracted RNA was performed by measuring concentration in Nanodrop 800 Spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA).

First strand cDNA synthesis was carried out by using 1 µg of total RNA in 13.5 µl volume and Tetro Reverse Transcriptase system (Bioline, London, UK). Initially, RNA was incubated at 70° C for 5 minutes in order to denature any secondary structure followed by incubating in ice for 5 minutes so that secondary structures cannot be formed again. Then 4 µl of 1x buffer, 1 µl of Oligo dT, 1 µl of 10 mM dNTP and 0.5 µl of Tetro reverse transcriptase was added into each sample. In order to detect any potential DNA contamination, a negative control was also used by using sterile water instead of the reverse transcriptase enzyme. The reverse transcription was performed by incubating at 42° C for 30 minutes followed by inactivating the enzyme at 85° C for 5 minutes. The synthesized cDNA was diluted 1:5 for the purpose of using in PCR or Q-RT-PCR.

2.7 PCR

2.7.1 Standard PCR

Standard PCR was performed by using following reagents in a total volume of 50 µl - 10 µl of 5x reaction buffer, 2 µl of $MgCl_2$ (50 mM), 1 µl of dNTPs (10 mM), 1 µl of forward primer (10 µM), 1 µl of reverse primer (10 µM) and 0.2 µl of Mango Taq DNA polymerase (Bioline, Alexandria, NSW, Australia), 5 µl of 50ng/µl template DNA and relevant amount of sterile distilled water to final volume. PCR was run in a thermal cycler having heated lid in the following way: 94° C for 5 minutes, (94° C for 1 minute, annealing temperature of ~60° C for 45 sec., 72° C for 1 minute/Kb as per expected size of the PCR product) 35-40x, 72° C for 10 minutes, hold at 15° C.

2.7.2 Agarose gel electrophoresis and visualisation of DNA

In order to determine the size of the PCR product, it was run on agarose gel (molecular grade; Bioline, Alexandria, NSW, Australia) prepared in TAE buffer (40 mM Tris Acetate and 1mM EDTA) which also contained Gold View ™ Nucleic acid stain (Acridine orange; SBS Genetech Co. Ltd., Beijing, China). As a reference, Easy ladder I was used (Bioline, Alexandria, NSW, Australia). The gel was then subjected to UV light using the Molecular Imager® Gel Doc™ XR System (Biorad, NSW, Australia) in order to visualize the PCR product. To enable visualization of smaller PCR products, high resolution agar (Sigma-Aldrich, NSW, Australia) was used.

2.7.3 PCR product purification

For cloning or sequencing, PCR products were purified by using Promega Wizard SV Gel and PCR Clean-Up System (Promega, Madison, WI, USA) and manufacturer's recommendations were followed in this regard. After final elution in sterile distilled water, concentration of the PCR product was estimated in Nanodrop 800 Spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA) and diluted to 50 ng/µl concentration for sequencing.

2.8 Cloning

2.8.1 Regular cloning

Purified PCR products were initially ligated to pGEM®-T Easy vector (Promega, Madison, WI, USA) by following the protocol provided by the manufacturer. Ligation reaction was used afterwards for transformation of α-Select chemically competent *E. coli* cells provided by

Bioline (Taunton, MA, USA). To this end, 5 µl of ligation reaction was mixed with 50 µl bacterial cell suspension and incubated on ice for 30 minutes. Then, the mixture was placed in a heat block at 42° C for 45 seconds followed by immediate transfer to ice for 2 minutes. The transformation reaction was then diluted by adding 400 µl of Luria broth medium and placed in a shaker for 1 hour at 37° C, 200 rpm. 225 µl of this reaction was then spread under sterile condition using a sterile glass spreader unto pre-poured LB agar medium plates which have 100 µg/ml ampicillin and 1 µl/ml X-gal. These plates were then incubated at 37° C overnight.

2.8.2 Colony PCR

From the blue/white colonies, certain number of white colonies from each genotype were picked and mixed with 5 µl of sterile water. This cell suspension was then incubated at 95° C for 5 minutes in order to disrupt the cell wall of the bacterial cells. After that, the mixture was used for the confirmation of expected inserts using both gene specific and vector specific primers present in a 20 µl of PCR master mix. Primers for the genes and vectors are given in table A1.3.

2.8.3 Plasmid preparation for sequencing

Colonies that were confirmed to have the proper insert via colony PCR were transferred to 5 ml of terrific broth medium (added 5 µl of ampicillin) and incubated at 200 rpm, 37° C overnight. Plasmid DNA from these broth culture were extracted according to manufacturer's instructions mentioned in Wizard plus SV miniprep DNA purification system provided by Promega. Concentration of the extracted plasmid was measured in Nanodrop 800 Spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA) and diluted to 100 ng/µl concentration for sequencing.

2.9 Sequencing facility and analysis of sequence

After purification, PCR products were sent for sequencing to Macrogen Inc. (Seoul, South Korea). Received sequences were trimmed at 5' and 3' end in order to remove sequences of poor quality using Geneious 8.0.4 software (Kearse et al. 2012). Manual editing of the trimmed sequences were performed for fixing bases which were identified wrongly. Finally, analysed sequences were annotated.

2.10 Molecular marker development for genotyping and mapping

2.10.1 Molecular marker design

Marker development strategy in the current study relied upon sequencing intronic regions from the parental genotypes as introns (Côrte-Real et al. 1994; Deulvot et al. 2010) are source of high genetic variation. If any size variation was revealed through visualization in PCR products obtained from the parental genotypes, then they were used directly for designing size markers. When such size variation was not gained, then the PCR products were purfied and sequenced. If the sequenced PCR product had polymorphism between the parental genotypes, then high resolution melting curve markers were generated where feasible. Otherwise, cleaved amplified polymorphic sequence (CAPS) or derived cleaved amplified polymorphic sequence (dCAPS) markers were developed. Information on the primers used for marker development are provided in table A1.3.

2.10.2 Marker gene selection for mapping

When possible, already available molecular markers from previous relevant publications (Aubert et al. 2006; Bordat et al. 2011) for same mapping population were used initially by Hecht et. al (unpublished data) and also during the course of current study. Genomic synteny between pea and *Medicago truncatula* (Kaló et al. 2004) was used to fine map *LATE3* and *LATE4* locus. In such case, *M. truncatula* genes that were located in the syntenic region between the flanking markers of corresponding pea linkage groups were used for further fine mapping. *M. truncatula* Mt4.0v1 database (https://phytozome.jgi.doe.gov/pz/portal.html#!info?alias=Org_Mtruncatula) was used for obtaining the genomic and transcriptomic sequences of *M. truncatula* genes within the aforementioned region. Orthologus pea transcriptome sequence was obtained via BLASTn search of the *M. truncatula* sequence against the pea RNA sequence gene atlas (http://bios.dijon.inra.fr/FATAL/cgi/pscam.cgi) database. Reciprocal BLAST was used in order to confirm presence of single copy of the *M. truncatula* gene. Due to unpredictability of unkown intron size in pea, PCR fragment generation was largely restricted to one intron only. Besides, *M. truncatula* genes having no intron within the region of interest were elimited for the marker design purpose.

Since the initially identified location for *LATE4* locus by Hecht et. al in pea *LGV* has poor synteny with *Medicago*, so the order of pea transcripts and their corresponding *M. truncatula* orthologoues mentioned in the consensus genetic map of *PsLGV* by Tayeh et. al 2015 was used for further marker development. However, the actual markers from Tayeh et. al 2015 were not exploited in the current study as the mapping populations used between these two studies were different.

2.10.3 Size markers

Size markers were developed usually by choosing PCR products that varied in 15 bp or more between the parental genotypes (≥10% of the length of PCR product) during the initial visualization of PCR products. New primers were developed proximal to the region of polymorphism for such PCR based size markers in order to increase comparative size differences and they were scored in the mapping population via standard PCR and visualization conditions. In the cases where sequenced PCR product revealed size differences between the parental genotypes, but were not suitable for directly using as a standard PCR based size marker, attempts were made to convert them into HRM markers initially. However, in some cases such size difference did not lead to generation of successful HRM marker because the polymorphic site had varying lengths of short tandem repeats. Under such situations, the polymorphic site (4-8 bp difference) was still used as a size marker by standard PCR where the PCR product was visualized after running for long time (>2 hour) in high resolution (3-4%) agarose gel.

2.10.4 High resolution melt (HRM) markers

High resolution melt (HRM) marker generation process involves distinguishing different genotypes in real-time depending on the temperature required to separate double stranded DNA due to the presence of polymorphic site(s) between those genotypes. HRM markers were designed around sites of insertion, deletion or SNP with fragment size limited to <200 bp. Such markers were initially tested on the parental genotypes and then scored in the respective mapping population using Rotorgene Q HRM machine (Qiagen). Samples for HRM run were prepared using CAS robotics 4 version 4.9.8 (1.6.61) software from a Corbettt Robotics CAS-1200™ pipetting robot (Corbett research, Australia). Reaction preparation was as follows for testing parental genotypes: 7.5 µl HMR PCR master mix provided by HMR PCR kit (Qiagen), 1.05 µl forward primer, 1.05 µl reverse primer, 2 µl template DNA and 3.4 µl

sterile mili-Q water. For scoring the marker following reaction was prepared: 5 µl HMR PCR master mix provided by HMR PCR kit (Qiagen), 0.7 µl forward primer, 0.7 µl reverse primer, 2 µl template DNA and 1.6 µl sterile mili-Q water. HRM run condition was set as follows: 95° C- 5 minutes, 50X (95° C- 5 seconds -> ~60° C annealing temperature-10 seconds ->72° C-20 seconds), 95° C -5 minutes (dissociation), 50° C – 5 minute (re-annealing), HRM: 60-90° C with temperature rising gradually by 0.1° C. Results from the HRM run were analysed using Rotor-gene screen clust HRM software (Qiagen).

2.10.5 Cleaved amplified polymorphic sequence (CAPS) markers

In some of the cases, the parental genotypes had a single nucleotide polymorphism (SNP) between them which were not suitable for getting a good HRM marker. If such SNP caused modulation of a restriction enzyme recognition site as determined by dCAPS finder 2.0 (http://helix.wustl.edu/dcaps/dcaps.html, (Neff et al. 2002) which could potentially generate clearly visible PCR products, then such SNPs were exploited for designing CAPS markers. Where applicable, new primers were generated in close proximity of the SNP to obtain a suitable PCR product for such CAPS markers. PCR products from the parental genotypes were initially subjected to corresponding restriction enzyme digestion according to manufacturer's instructions (New England Biolabs Inc., Ipswich, MA). Successfully digested markers were then scored in the respective mapping populations.

2.10.6 Derived cleaved amplified polymorphic sequence (dCAPS) markers

dCAPS markers were designed using dCAPS finder 2.0 (http://helix.wustl.edu/dcaps/dcaps.html, (Neff et al. 2002) via introduction of one mismatch in the primers for those polymorphic SNPs where alteration in restriction enzyme recognition site did not occur between the parental genotypes or no cost effective restriction enzyme was avaiable for the CAPS marker. Testing and scoring of dCAPS marker were similar to the methodology used for CAPS markers.

2.10 Generation of genetic map

Linkage groups were generated by calculation of genetic distance in the form of recombination frequency (known as centi Morgan, cM) between molecular/DNA and phenotypic markers by exploitation of segregation data. The software JoinMAP 4 (Kyazma

B.V., Wageningen, Netherlands) was used for this purpose by using a threshold of LOD value of 3.0.

2.11 Statistical analysis

Where applicable, One-way ANOVA followed by Dunnett test was performed to determine the level of significance between the mean of WT and various mutant genotypes using Prism software. p-value significance level used was ****≤ 0.0001, ***≤ 0.001, ** ≤0.01, *≤ 0.05, ns > 0.05.

Chapter 3: Phenotypic characterization of mutants for *LATE BLOOMER 3* (*LATE3*) and *LATE BLOOMER 4* (*LATE4*) loci

3.1 Introduction

The regulatory process that mediates flowering in *A. thaliana* as well as identification of different orthologues in pea have been mentioned in chapter 1. The process of identifying developmental genes in pea have largely followed the classical forward genetics approach which is based on the principle of using phenotypic variation of a mutant to identify the gene and reveal its function (Weller et al. 2013). As part of this, induced mutants are generated by using chemical mutagens such as ethyl methanesulfonate (EMS) since T-DNA or transposon based insertional mutation platform is yet to be established in pea. One advantage of using chemical mutagens is that it does not require to usage of complicated facilities. Usually the seeds are soaked to a certain concentration of EMS for inducing random mutations. Phenotypic characterization of novel mutants of interest with relevance to flowering assist in understanding the genetic basis of the observed phenotypic alteration. Identifying new flowering genes in a model legume such as pea is of high importance in order to deal with challenges of food and fooder security (Weller et al. 2013).

An EMS mutant secreeing previously led to the identification of two extremely late flowering mutants namely *late3* and *late4* (Weller JL 2007). These were generated in the background of wild-type pea line NGB5839. The NGB5839 line has a mutation (*le-3*) in GA 3β-hydroxylase gene *LE* which is defective in bioactive GA, GA1 (Lester et al. 1999). Phenotypic comparison with classical pea cultivar “Torsdag” established the fact that mutation in *LE* gene did not significantly alter node of flowering initiation and other flowering related traits under both long day (LD) and short day (SD) conditions in NGB5839 (Hecht et al. 2007). Therefore, NGB5839 has been wildely used as the wild-type genotype in investigating flowering time mechanism in pea by Weller et al. at the School of Natural Sciences, University of Tasmania in various studies specifically due to its small size and easy handling purposes. Other advantages of using this genotype is the simplicity in generating isogenic double mutants as well as the availability of relevant molecular tools such as genetic maps and markers.

A major goal of EMS induced mutant screening is to identify functional alleles of the same gene (Weller et al. 2013). This is very useful because phenotypic and molecular characterization of multiple alleles of the same gene is likely to eliminate the necessity of conducting reverse genetics based experimentations later on. Allelism test is performed by crossing two recessive mutants of interest (Peters et al. 2003). If the observed phenotype of the resultant heterozygote is wild type, then it can be assumed that the two mutations complement each other and they are located in two different genes. In contrast, generation of mutant phenotype in the heterozygote would mean that they do not complement each other and the mutations are positioned in the same gene. Such an allelism test identified three recessive mutant alleles known as *late3-1, late3-2* and *late3-3* for *LATE3* locus and two recessive mutant alleles namely *late4-1* and *late4-2* for *LATE4* locus (Weller 2007). Under LD conditions, *late3* and *late4* mutants showed significantly delayed flowering compared to wild type NGB5839 giving hints that both the underlying genes are important for general promotion of flowering in pea. Flowering occurs in these mutants after node 35 compared to node 16 for NGB5839 under LD conditions. Besides, reproductive structures in the initial reproductive nodes are sterile and floral reversion occurs after a period of reproductive growth. Few reproductive pods grow successfully, but those are weak and generate very low number of seeds. Under cold environment, these mutants can grow indefinitely and have a prolonged reproductive phase. Increased aerial branching at a later stage of development is another distinct feature of these mutants. Due to their phenotyping difference with the wild type in flowering and other important traits such as internode length, aerial branching, floral morphology, duration of reproductive stage, these mutants are of great interest for functional characterization studies in order to reveal the identity and function of the underlying genes that seem to be crucial for various reproductive and vegetative traits.

3.2 Chapter aims

In general, the goal of this chapter was to investigate the phenotypes of *late3 and late4* mutants in detail. As part of that, various flowering and reproduction related traits from both *late3* and *late4* mutants were studied comprehensively with higher priroity. Further phenotypic characterization involved studies on various traits of leaf, stem, root as well as

overall architecture of *late3* and *late4* mutants. To this end, different traits from early to late growth stage were inspected carefully.

3.3 Materials and methods

Various materials and methods used for this chapter have been mentioned in general materials and methods of chapter 2.

3.4 Results

3.4.1 LATE3 and LATE4 affect flowering time

In preliminary studies, *late3* and *late4* mutants exhibited extremely delayed flowering phenotype compared to the wild type (WT) NGB5839 when grown under long day (LD) condition (Weller et al., unpublished). Flowering time of the various *late3* and *late4* mutant alleles under LD was also determined in the current study. Figure 3.1 shows that flowering commenced in WT at node 18 while the process started at node 28 in *late3-1* and at around node 42 in the stronger alleles *late3-2* and *late3-3* (Figure 3.1 A, B, D). Likewise, flowering was

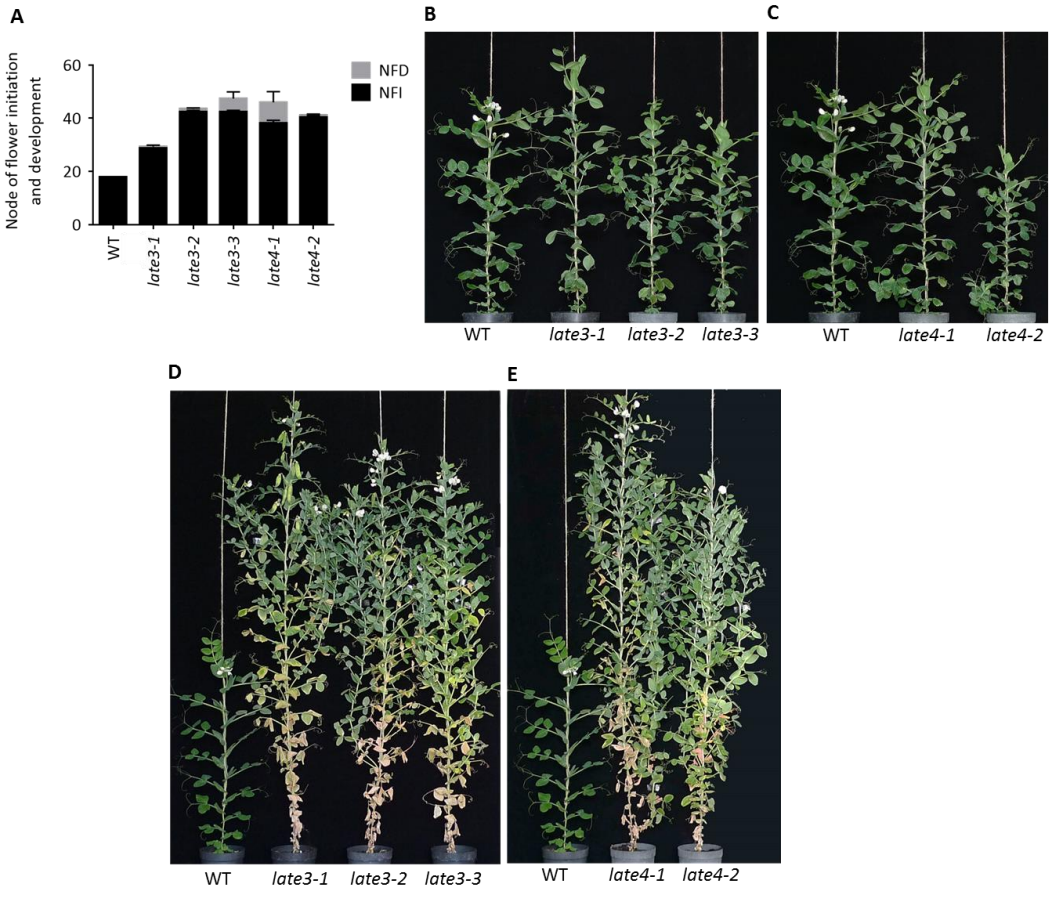


Figure 3.1. Flowering initiation in wild type (WT) NGB5839 and *late3* and *late4* mutants grown under LD. **(A)** Node of flower initiation (NFI) represent the appearance of first flower bud along the main stem whereas node of flower development (NFD) represent the first fully developed flower in WT, *late3* and *late4* mutants, data represents mean ± SE for n = 6-10 plants. **(B-C)** Representative 75 day old WT and *late3* and *late4* mutants. **(D-E)** Representative 130 day old *late3* and *late4* mutants plants grown under LD, WT plant (62 day old) shown were sown during a different time point and it is used in order to compare node of flower initiation in WT and *late3* and *late4* mutants.

initiated in *late4-1* and *late4-2* at node 38 and 40 respectively (Figure 3.1 A, C, E). Delay in floral initiation for all these mutants in comparison to WT was statistically significant ($p < 0.0001$). In case of WT, *late3-1*, *late3-2* and *late4-2* genotypes the first flower transformed into a fully open flower (Figure 3.1 A). In contrast, this process was delayed in *late3-3* and *late4-1* mutants by nearly five and eight nodes respectively as some of the initial floral buds in these mutants were malformed and failed to open. These phenotypic results indicate that *late3-1* is the weaker allele as it showed less severe phenotypes compared to the other mutant alleles. Overall, the findings suggest that *LATE3* and *LATE4* are essential for general promotion of flowering.

Leaf formation rate of wild type (WT) NGB5839 as well as *late3* and *late4* mutants were also examined in this study because some of the flowering mutants previously showed marked difference in leaf formation rate compared to WT (Weller et at., unpublished data) (Figure 3.2 A – B). Leaf emergence rate/week was similar between WT and all the alleles of *late3* and *late4* mutants amounting to approximately 18 until the appearance of first flower in the WT.

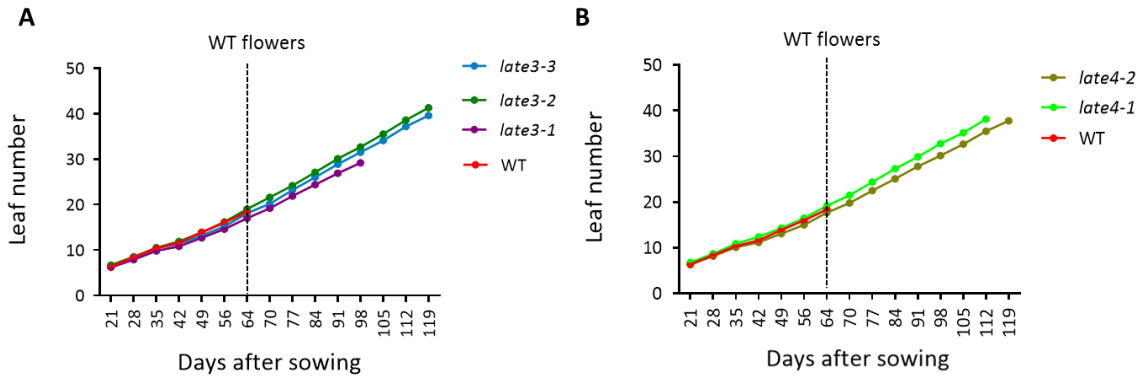


Figure 3.2. (A-B) Comparison of various leaf growth rate between wild type (WT) NGB5839 and *late3* and *late4* mutants grown under LD. Data represents mean ± SE for n = 6-10 plants, leaves were counted each week starting from day 21 until flowering was initiated in each genotype, dashed line indicates initiation of flowering in WT.

However, slight reduction was observed between the stronger and weaker alleles of *late3* from the period of 84 days after sowing until the initiation of flowering in the weaker *late3-1* mutant allele. Similar type of alteration also occurred between the two *late4* alleles from 84 days after sowing until initiation of flowering in weaker *late4-1* mutant allele. Nevertheless,

these results implicate that the underlying *LATE3* and *LATE4* genes are not crucial in determining leaf formation rate.

3.4.2 LATE3 and LATE4 act in determining inflorescence development and floral organ identity

Proper inflorescence development is important in plants for downstream traits such as yield. Various pea mutants with abnormalities in inflorescence architecture have been characterized in previous studies such as *uni* (Hofer et al. 1997), *pim* (Taylor et al. 2002), *veg1* (Berbel et al. 2012) and *veg2* (Sussmilch et al. 2015). In an initial study (Weller et al., unpublished) atypical patterns in inflorescence development have been observed in *late3* and *late4* mutants, however this was not investigated closely. Therefore, the current study shed light upon understanding potential role of *LATE3* and *LATE4* genes in inflorescence generation by studying phenotypic variation of this particular plant organ in the respective mutants.

It was observed that both *late3* and *late4* mutants generated various types of aberrant inflorescences which did not resemble the known morphology of pea inflorescence (Figure 3.3, A-C). Pea in known to possess a compound inflorescence where two floral meristems (FM) are usually generated from secondary inflorescence meristems (I_2) rather than primary inflorescence (I_1) which is the case in *A. thaliana* (Benlloch et al. 2007; Benlloch et al. 2015). Thus, two flowers are produced mostly in an inflorescence. The observed structure of almost all WT inflorescences from this study resembled this aforementioned outlook (Figure 3.3 A-C). In contrast, *late3* and *late4* mutants had inflorescences which were similar to WT as well as various uncommon inflorescences where regular floral organs did not form or were replaced by vegetative organs (Figure 3.3 A-C). Such kind of abnormalities were more severe in the stronger alleles of *late3*, i.e., *late3-2* and *late3-3* than *late3-1.* The types of irregular inflorescences generated in the *late3* as well as in *late4* mutants included carrying small floral buds either single or in pair having no sepals and petals which senesced quickly and did not blossom, flowers having irregular pattern in sepal and petal or being replaced entirely by leaf like structure. Peduncle length of these malformed inflorescences also varied widely in comparison to WT. These results lead to speculation that *LATE3* and *LATE4* genes are crucial for regulating proper inflorescence architecture.

Figure 3.3. (**A-C)** Different types of abnormal inflorescences carrying unusually structured flowers in *late3* and *late4* mutants, examples from *late3-2*, *late3-3* and *late4-2* are illustrated here. Scale bar represents 10 mm.

WT pea flowers are zygomorphic (Wang et al. 2008; Sussmilch 2014) and they carry three different types of petals (Figure 3.4 A-D). Petals are five in number which are made up of one standard, two wings and two adjoined keel petals. Likewise, there are five sepals comprising of two adaxial and three abaxial all of which are fused at the base. Moreover, they have 10 stamens being merged with each other and a single carpel. In order to get a deeper insight into the inner structure of the abnormal flowers observed in both *late3* and *late4* mutants during this study, one such flower was dissected from the *late4-2* mutant (Figure 3.4 E-G). As was observed, this flower undergone modulation in petal formation as it did not develop any standard and keel petals. Besides, it also demonstrated alteration in sepal, i.e., carrying seven irregular sepals. Similarly, stamen generation was affected in this flower where seven stamens were formed instead of regular 10. As it was not the primary focus of the current study to investigate in detail about the internal structure of various types of abnormal flowers

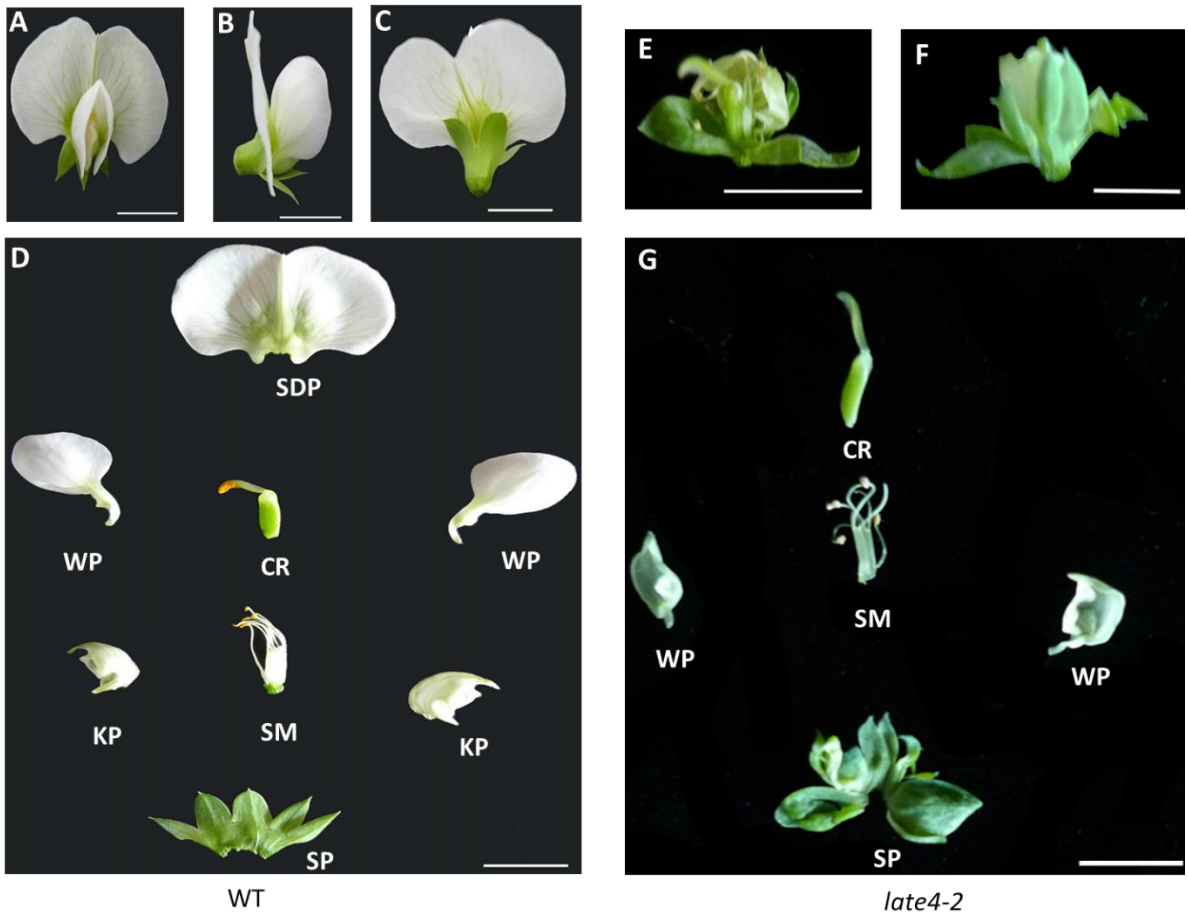


Figure 3.4. Comparison of morphology of a standard flower generated in wild type (WT) NGB5839 and a malformed flower produced in *late4-2* mutant grown under LD condition.

(A-D) WT flower; **(A)** Front view of WT flower; **(B)** Side view of WT flower; **(C)** Top view of WT flower; **(D)** Various parts of a dissected WT flower carrying five petals – one standard petal (SDP), two wing petals (WP), two keel petals (KP), five sepals (SP), 10 stamens (SM) attached at base and a central carpel (CR).

(E-G) *late4-2* flower; **(E)** Front view of malformed *late4-2* flower; **(F)** Top view of malformed *late4-2* flower; **(G)** Components of malformed *late4-2* flower - two wings petal (WP), seven sepals (SP), seven stamens (SM) and a central carpel (CR). Scale bar represents 10 mm.

in *late3* and *late4* mutants, so such activities were not performed within the limited time of this study. Nevertheless, the observed abnormalities in the dissected flower of *late4-2* and presence of other types of malformed flowers in *late3-2* and *late3-3* give hints that the underlying genes play a major role in establishing specificity of various floral organs.

3.4.3 LATE3 and LATE4 are required for appropriate post-flowering reproductive development

Duration of reproductive phase and entire growth period are important in plants since these are related to yield and timing of senescence respectively. In pea, this is routinely performed by counting the number of reproductive as well as total nodes (Weller et al. 2004; Hecht et al. 2007; Liew et al. 2009; Liew et al. 2014; Sussmilch et al. 2015; Ridge et al. 2016; Rubenach et al. 2017). Therefore, data was collected in the present study for the aforementioned traits in order to reveal whether both these processes are affected by mutation in *LATE3* and *LATE4* genes.

In many plants such as *A. thaliana*, a transient, inductive floral promoting signal is needed for making stable transition from vegetative to reproductive development (Corbesier et al. 2007). The establishment of flowering is not sufficient for successive reproductive growth, rather it needs to be maintained (known as floral commitment) specially when the primary inductive signal is no longer present (Adrian et al. 2009). In some *A. thaliana* mutants as well as other species, continuous flow of floral stimulus is needed for floral commitment and anomaly in this process leads to both inflorescence and floral reversion (Tooke et al. 2005; Adrian et al. 2009). In case of former situation, the shoot stops producing flowers and reverts to formation of leaves whereas the latter situation occurs when a flower starts to produce vegetative organs instead of floral organs. This latter scenario suggests that both establishment and preservation of floral meristem identity is crucial for reproductive development in plants (Parcy et al. 2002; Tooke et al. 2005; Adrian et al. 2009). In a preliminary study, inflorescence reversion in *late3* and *late4* mutants were observed (Weller et al., unpublished) which gave hints that the underlying genes are important for sound maintenance of reproductive development, but it was not cautiously studied. Hence, the current study included careful observation of this trait as well.

The WT generated around seven reproductive nodes once flowering was commenced at node 18 under LD condition as they went on to grow until node 25 when the apex was senesced (Figure 3.5 A). In previous studies, more or less same numbers of TN and RN was observed in WT (Sussmilch et al. 2015; Ridge et al. 2016; Rubenach et al. 2017). In contrast, all the mutant alleles had significantly higher numbers of initial reproductive nodes (I-RN) ($p < 0.0001$) as well as total nodes (TN) (Figure 3.5 A). Among the *late3* mutants, stronger alleles *late3-2* and

late3-3 had higher numbers of nodes for both the traits compared to weaker *late3-1* allele. Besides, *late4* alleles had more TN and I-RN compared to *late3* mutants. This number of I-RN ranged from 12-22 nodes for the different *late3* and *late4* mutants.

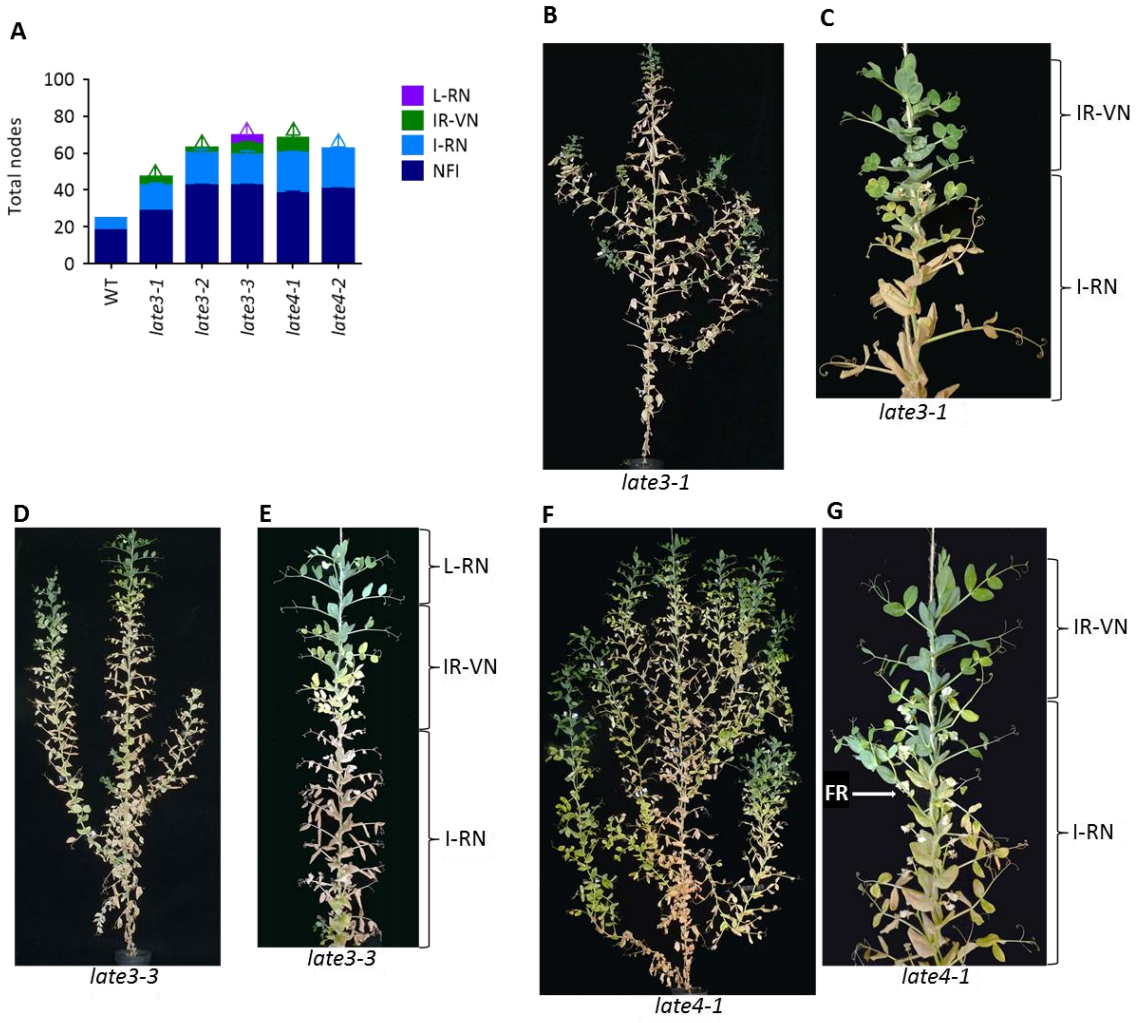


Figure 3.5. Inflorescence reversion at main stem of *late3 and late4* mutants grown under LD. NFI = Node of flowering initiation, I-RN = Initial reproductive nodes, IR-VN = Inflorescence reversed vegetative nodes, L-RN = later reproductive nodes, FR = Floral reversion. **(A)** Comparative graphs showing different types of node in WT and *late3* and *late4* mutants, data represents mean ± SE for n = 6 plants. Arrow indicates continuing growth. **(B-C)** *late3-1* mutant showing different types of nodes in main stem. **(D-E)** *late3-3* and **(F-G)** *late4-1* mutants showing various types of node and floral reversion in the main stem. All pictures were taken 180 days after sowing.

Observations also confirmed the occurrence of inflorescence reversion in all mutants apart from *late4-2* which have been demonstrated by formation of three to eight inflorescence reversed vegetative nodes (IR-VN) (Fig 3.5 A-G). Similar pattern of growth was observed in

many of the lateral branches as well. Interestingly, the strongest allele of *late3*, i. e., *late3-3* then recommenced forming flowering nodes generating what is termed as later reproductive nodes (L-RN) (Figure 3.5 A, D, E). In addition to this reversion of the primary inflorescence (I_1), the *late3* and *late4* mutants also displayed floral reversion (I_2) where vegetative shoot emerged from the centre of inflorescence (Fig 3.5 G). Since the current study had to be ceased 180 days after sowing due to time limitation and emergence of pests, so it was not possible to determine the entire growth period L-RN for all mutant genotypes apart from one such stage in *late3-3*. It is very unusual for pea mutants to grow for such long period and no other mutants previously showed such prolonged growth. Nevertheless, the results obtained within the duration of this study demonstrate that *LATE3* and *LATE4* are key for accelerating overall plant senescence and promote rapid completion of the reproductive growth stage. Moreover, it could be apprehended that they act in suppressing inflorescence and floral reversion, and thus they play role in steady perpetuation of inflorescence development and floral meristem identity once reproductive growth has been initiated.

3.4.4 Successful generation and structure of pods as well as seed development procedure is mediated by LATE3 and LATE4

Pods are important for protecting seeds from pests and pathogens (Bennett et al. 2011). Similarly, seed size and number are considered as the main components of yield specifically in crop plants, therefore such traits are crucial from evolutionary and applied perspective (Gnan et al. 2014). As it was found that mutation in *LATE3* and *LATE4* genes affect various reproductive development involving flowers, so it looked relevant to investigate whether the downstream reproductive processes, i.e., pod and seed development are also affected by these mutations. Due to the aforementioned importance of pods and seeds in plants, this study carefully characterized different phenotypic traits in these two plant organs. As the WT had a shorter reproductive stage, so it produced significantly lower number of flowers than the various mutant alleles of *late3* and *late4* (Figure 3.6 A, $p < 0.0001$). Even though *late3* and *late4* mutants have produced more flowers than WT, only a small percentage of these flowers were transformed into a seed carrying pod (Figure 3.6 B). In WT, nearly 65% of the total flowers led to generation of pods carrying atleast one seed (Figure 3.6 B). In contrast, this conversion rate was substantially lower in *late3-1* at 35% (Figure 3.6 B). The stronger alleles of *late3*, i.e., *late3-2* and *late3*-3 mutants had even lower transformation rate amounting to

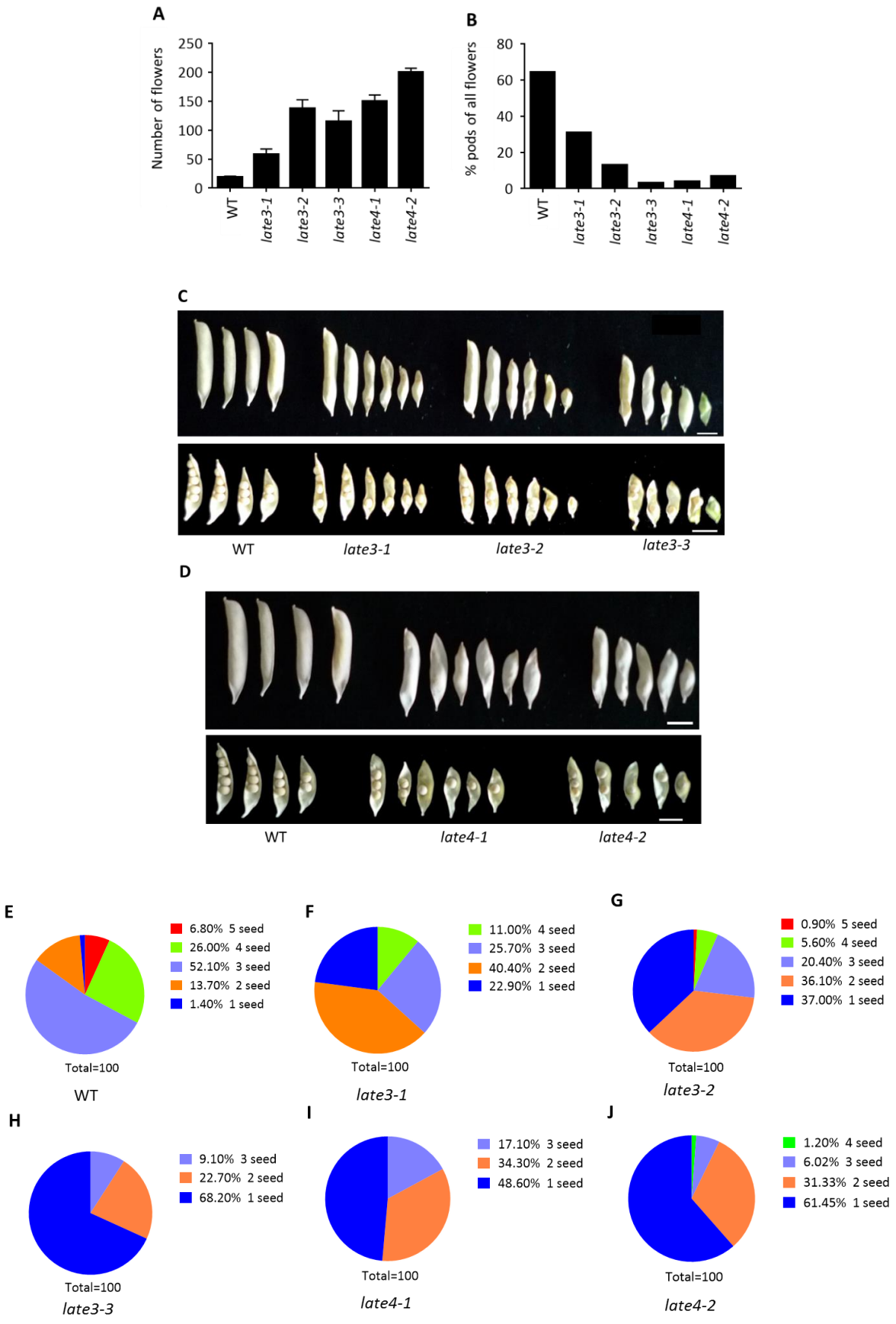


Figure 3.6. Variation in quantity and success rate in development of all and various types of pods in wild type (WT) NGB5839 and *late3* and *late4* mutants grown under LD. **(continued next page)**

Figure 3.6. (continued) (A) Comparison of the total number of flowers produced in wild type (WT) NGB5839 and different mutant alleles of *late3* and *late4*, data represents mean ± SE for n = 6 plants. **(B)** Percentage of seed bearing pods generated out of total number of flowers in WT and *late3* and *late4* mutants. Data of total number of flowers and seed carrying pods in main and lateral branches of WT and mutants were collected at the time of harvest (180 days after sowing) and calculated as a mean ± SE for n = 6 plants. (**C-D)** Comparison of morphological structure of different types of pods (unopen and open) carrying varying number of seeds in WT and *late3* and *late4* mutants, scale bar represents 10 mm. **(E-J)** Percentage of pods producing five, four, three, two and one seed in WT and *late3* and *late4* mutants.

only 10% and 5% respectively (Figure 3.6 B). The proportion of seed bearing pods for both the alleles of *late4* was below 10% (Figure 3.6 B).

Majority of pods in the WT were more or less uniform in size and shape while *late3-1* mutant showed marked variation compared to the WT even though they generated some pods having similar structure as the WT (Figure 3.6 C). Other types of pods in *late3-1* were smaller in length in comparison to WT. *late3-2* developed more of the pods that were shorter compared to WT (Figure 3.6 C). *late3-3* had the strongest phenotype as it did not generate any pods that is similar to the WT, rather all the pods were from smaller size category (Figure 3.6 C). Similarly, *late4-1* and *late4-2* mostly generated pods which manifested distinctly reduced size compared to WT (Figure 3.6 D). Majority of these smaller pods from both the mutants had distorted shape as well. So, these results suggest that *LATE3* and *LATE4* regulate the transition from flower to pod formation which is probably related to floral fertility and they are also important for the adequate size and shape of the pods.

The pods were divided into five categories based on the number of seeds that they were bearing. These categories were pods carrying five, four, three, two and one seed. Majority of the pods from the WT genotype had three and four seeds amounting to nearly half and a quarter of all pods respectively (Figure 3.6 E). The percentage of two and five seed carrying pods were 13.7 and 6.8 respectively. The rest were one seed bearing pods. The *late3-1* mutant did not produce any five seed bearing pods (Figure 3.6 F). While the majority of the pods in *late3-1* had two seeds (40.40 %), nearly a quarter each of the pods were holding three and one seed. The effect was even severe in the stronger *late3-2* and *late3-3* mutants which mostly produced pods carrying two and one seed (Figure 3.6 G-H). Situation for *late4-1* and *late4-2* was more or less same to what was observed for *late3-2* and *late3-3* where most of the pods had only two and one seed (Figure 3.6 I-J). Overall, the stronger the mutant allele

was, the higher was the proportion of pods carrying low number of seeds. Thus, it is evident that *LATE3* and *LATE4* play strong role in preventing seed abortion and thereby regulate total yield.

As the dry weight of the seeds were measured, it was found that WT had significantly higher weight than all the mutant alleles of *late3* and *late4* (Figure 3.7 A, $p < 0.05$). Besides, *late3-1*

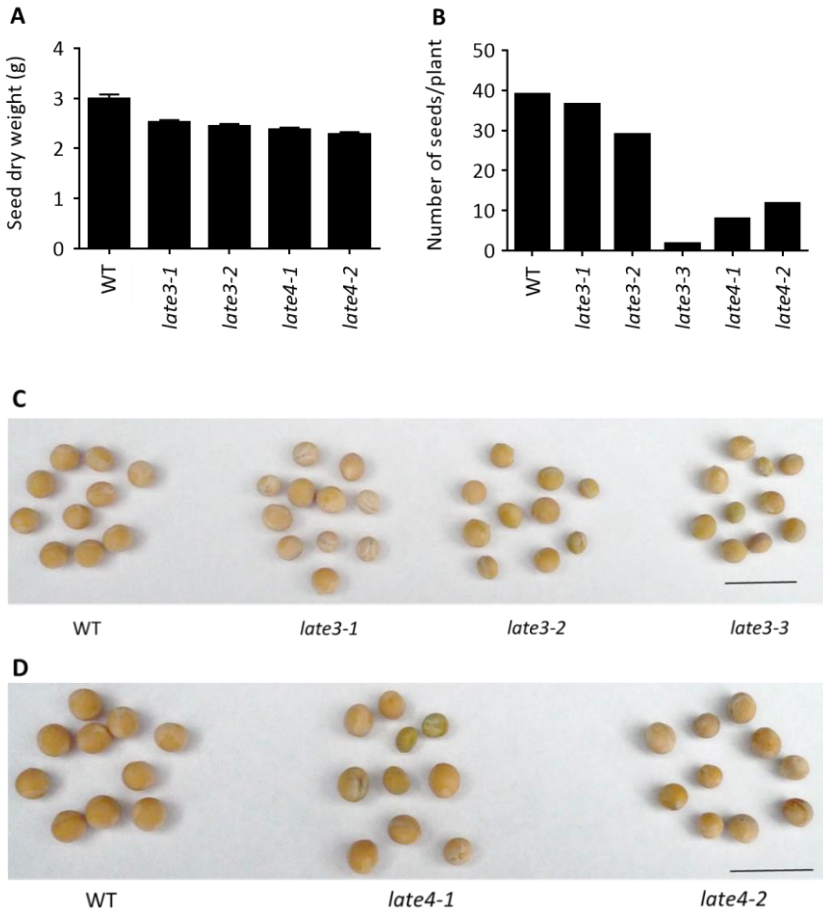


Figure 3.7. Various yield related characteristics of seeds between wild type (WT) NGB5839 and *late3* and *late4* mutants grown under LD. **(A)** Seed dry weight. Data represents mean ± SE for n = 3 independent measurements each of which consisted of 10 seeds. **(B)** Average number of seeds produced in a single plant. Seeds were collected in bulk from n = 6 plants at the time of harvest and then mean value was calculated. **(C-D)** Comparison of seed size and shape of WT and *late3* and *late4* mutants. Scale bar represents 10 mm.

generated almost similar number of seeds/plant like WT which was just below 40 while *late3-2* yielded nearly 10 seeds less (Figure 3.7 B). The number of seeds for each plant of the *late3-3* and the two mutant alleles of *late4* were considerably less compared to WT which was

below 10 (Figure 3.7 B). Seeds generated in WT were more or less equal in size and shape whereas all the mutant alleles of *late3* and *late4* generated seeds of varying size which were mostly smaller than the WT (Figure 3.7 C-D). These results provide hints that *LATE3* and *LATE4* have effect on seed size and shape.

3.4.5 LATE3 and LATE4 negatively regulate branching

It was previously reported in a study involving various pea cultivars and branching mutants that five different branching patterns occur in pea (Arumingtyas et al 1992), i. plants with no branches, ii. aerial branching at upper nodes only, iii. basal branching at basal nodes only, iv. upper and lower node branching with a gap in between and v. branching at all or nearly all vegetative nodes. In another study, removal of flowers from the wild-type plants resulted in increased branching (Lockhart and Gottschall 1961) which leads to understanding that delayed appearance of flowering and pods in late flowering mutants could be correlated to branch outgrowth. Previously, various photoperiod-sensitive late flowering pea mutants such as *late1*, *late2* and *phyA* mutants exhibited increased basal branching similar to SD-grown WT plants (Hecht et al. 2007). In contrast, other late flowering mutants which do not affect photoperiod sensitivity such as *gigas* (Hecht et al. 2007), *veg1* (Berbel et al. 2012), *veg2* (Sussmilch et al. 2015) showed aerial branching rather than basal branching. Branching in *late3* and *late4* mutants have been observed previously (Weller et al., unpublished data), but it was not studied closely. Therefore, the current study involved careful investigation on the patterns of branching in these two mutants due to relevance of this trait to flowering time.

As was observed, total branching (Figure 3.8 A, $p < 0.0001$) was significantly higher in *late3* and *late4* mutants compared to WT. All the three mutant alleles of *late3* had aerial branches only (Figure 3.8 B). For individual *late3-2* and *late3-3* plants, branching started at around 11-12 nodes which continued until around node 18 (Figure 3.8 B). It was then followed by a gap with branches re-appearing at around node 28, then a second gap occurred and branches again emerged at around 34-35 nodes which prevailed until the initiation of flowering. In case of *late3-1*, formation of first major branching took place at around node 16 for one to two successive nodes, then again at around node 20-21 and finally approximately at node 25 which continued until the penultimate node of floral initiation (Figure 3.8 B). Only *late3-2* mutants showed secondary branches in addition to primary branches. The length of branches

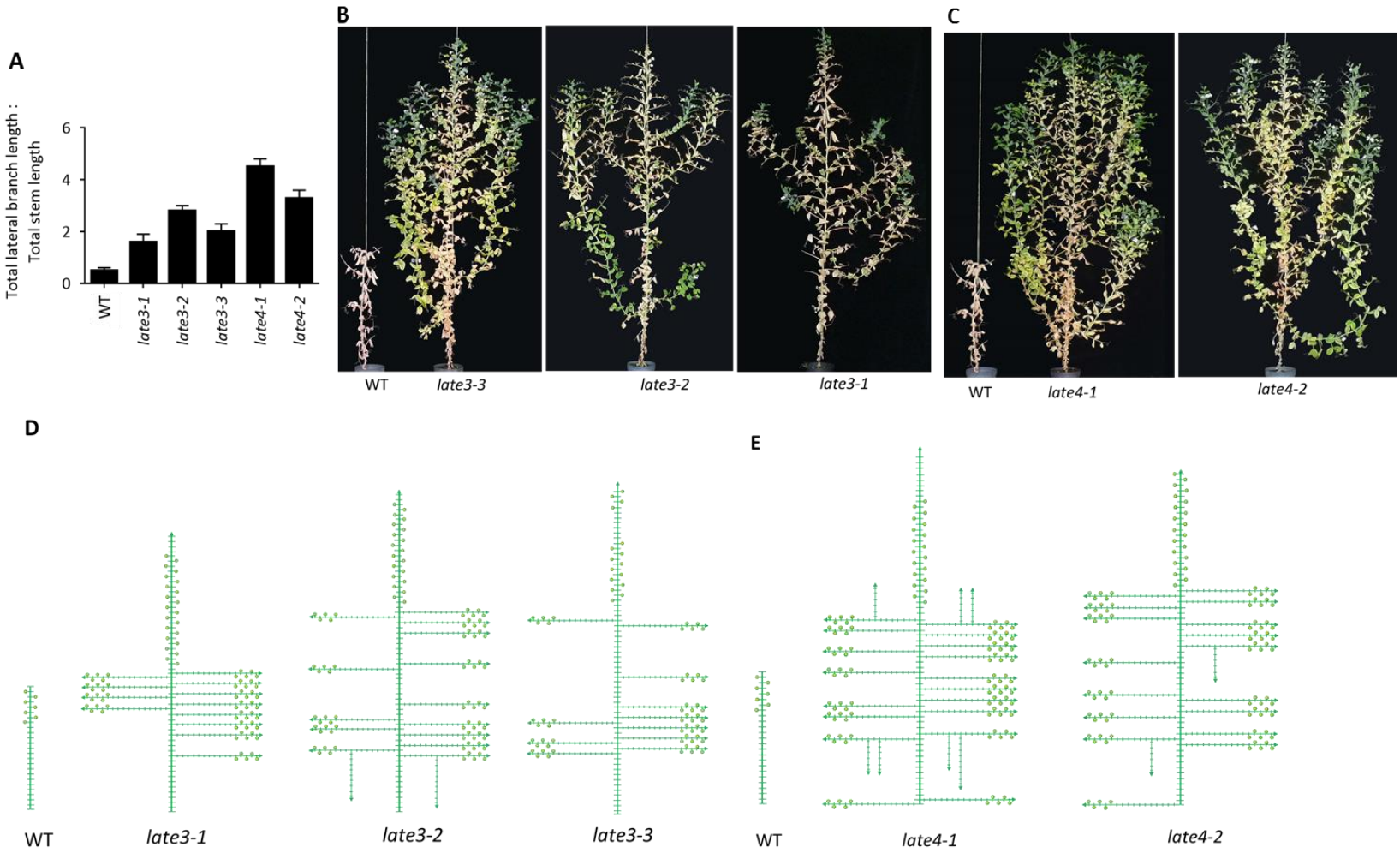


Figure 3.8. Branching in *late3* and *late4* mutants grown under LD. (**A**). Total branching as a ratio of total branch length and total stem length. Total length of all lateral branches with ≥ 5 cm and total stem length were measured at the time of harvest, i.e., 180 days after sowing Data represents mean ±SE 6 plants. **(continued next page)**

Figure 3.8 (continued) (B-C) Comparative picture of WT and *late3-3* and *late4* mutants 180 days after sowing. WT (105 days old) which stopped growing and senesced was not sown at the same time (included for comparison); *late3-2*, *late3-1* are in slightly different resolution as *late3-3*, WT; *late4-2* is also not in the same resolution as WT and *late4-1*, used here just for showing plant architecture at the time of harvest. **(D-E)** Schematic diagram drawn from six plants showing overall branching pattern observed at the time of harvest (180 days after sowing) in wild type (WT) NGB5839, *late3* and *late4* mutants. Long vertical line represents main stem, small horizontal lines in the main stem indicate a node where the line at the bottommost mark represent node 1, horizontal lines coming out from a node represent primary lateral branches while further branching in presented by vertical lines, star shaped green and yellow structure represent reproductive organ, single arrow indicates continuing growth.

was mostly longer in the stronger alleles, i.e., *late3-2* and *late3-3* in comparison to their weaker counterpart *late3-1*.

In case of individual plants of *late4-1* and *late4-2* (Figure 3.8 C), the tendency of aerial branch formation was more or less similar to that of *late3-2* and *late3-3* as all these mutants flowered nearly at same node. This means that *late4-1* and *late4-2* also had two gaps in branch formation. However, one striking feature for these two *late4* mutants was that they grew branches at basal nodes also which was not observed in any of the *late3* mutants. Besides, branches were longer in these mutants compared to *late3* mutants. In addition, both mutants developed secondary branches. Since individual plants (n=6) within a mutant genotype formed branches more or less in a similar pattern but at specifically different nodes, the overall appearance of branching at various nodes in the *late3* and *late4* is depicted in figure 3.8 D and 3.8 E respectively.

Taken together, results from the current study established role of *LATE3* and *LATE4* in prevention of branch formation in pea.

3.4.6 LATE3 and LATE4 are important regulator of leaf, stem and internode development

Leaf is the organ where photosynthesis takes place and therefore is the most vital plant component in establishing carbon economy (Edwards et al. 2014). It is the site which the plant uses in order to perceive the variation in environmental cues such as light quality, humidity, CO_2 concentration etc. on a constant basis. Hence, morphological appearance for various traits of leaf in the *late3* and *late4* mutants were carefully observed in this study which was not performed thoroughly in the preliminary study (Weller et al., unpublished data). Measurement of such traits was also an attempt to understand whether *LATE3* and *LATE4*

affect early stage growth characteristics or not. A typical pea leaf is pinnate and consisted of pairs of lateral organs having discrete identities where stipules, leaflets and trendrils are organized along petiole-proximal rachis in a descending order (Figure 3.9 A) (Tattersall et al. 2005; Demason and Chawla 2006). It was observed that *late3-2* had significantly smaller leaf area in comparison to WT (Figure 3.9 B, $p < 0.0001$). A similar pattern was observed for the *late4-1* mutant (Figure 3.9 B, $p < 0.0001$). Thus, *LATE3* and *LATE4* seem to positively regulate leaf size. Each of the mutant alleles of *late3* and *late4* had significantly shorter petiole and proximal rachis (Figure 3.9 C – D, $p < 0.0001$). In this case, stronger effect was also observed for these traits in the stronger alleles of both mutants. Altogether, these results suggest that *LATE3* and *LATE4* are essential for promotion of growth of organs that forms the entire structure of leaf.

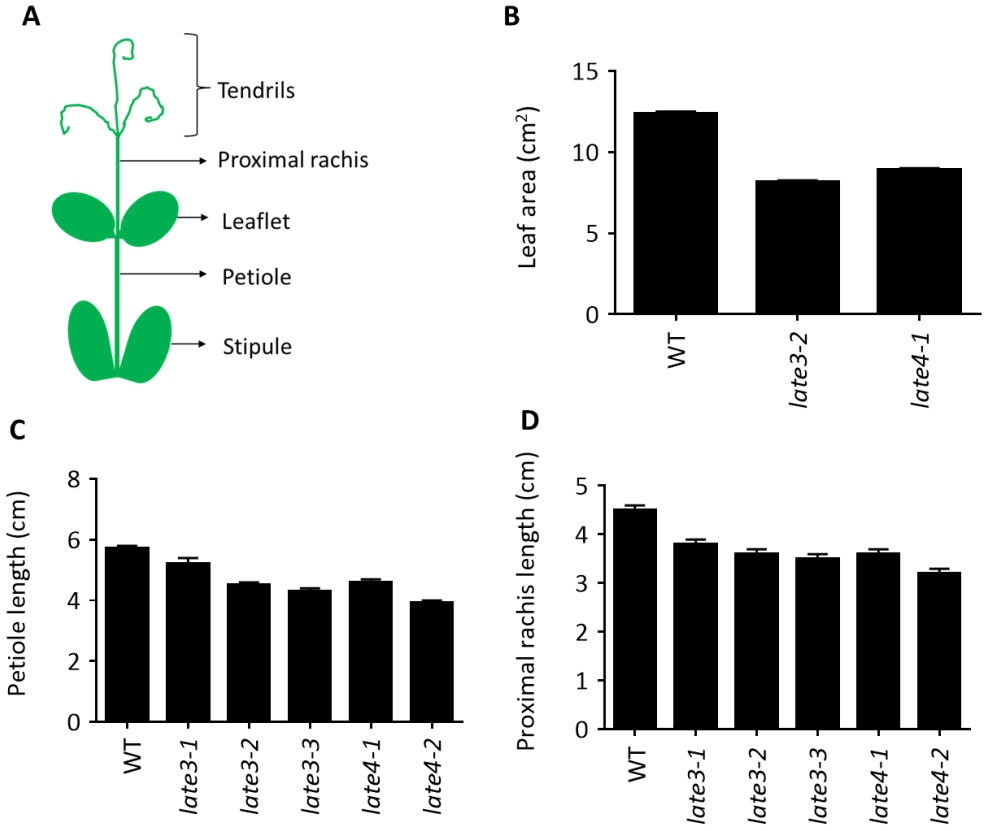


Figure 3.9. Comparison of various leaf morphological traits between wild type (WT) NGB5839 and *late3* and *late4* mutants grown under LD. **(A)** Schematic diagram showing various parts of a wild-type pea leaf. **(B)** Leaf area, data represents mean ± SE for n = 9-12 plants, leaf area were calculated for fully expanded single leaflet of 8th leaf. **(C)** Petiole length, data represents mean ± SE for n = 6-10 plants, measurements was made for petiole of 10th leaf **(D)** Proximal rachis length, data represents mean ± SE for n = 6-10 plants, measurements was made for proximal rachis of 10th leaf.

Stem acts as a bridge between root and leaves which assist in water, nutrient and photosynthate transportation. Besides, it is the carrier of all other above ground vegetative organs. As the current study focused on observation of alteration in phenotypic traits at different vegetative organs, so stem diameter of *late3* and *late4* mutants were compared with that of wild type NGB5839. Significant decrease in stem diameter was found for all the three mutant alleles of *late3* and two mutant alleles of *late4* relative to WT (Figure 3.10 A, $p < 0.0001$). Besides, the overall pattern in reduction was more pronounced in *late4* mutants than *late3* mutants. All the mutants had significantly longer stem compared to the WT (Figure 3.10 B, $p < 0.0001$). Based on these results, it can be stated that mutation in *LATE3* and *LATE4* genes affect proper development of stem.

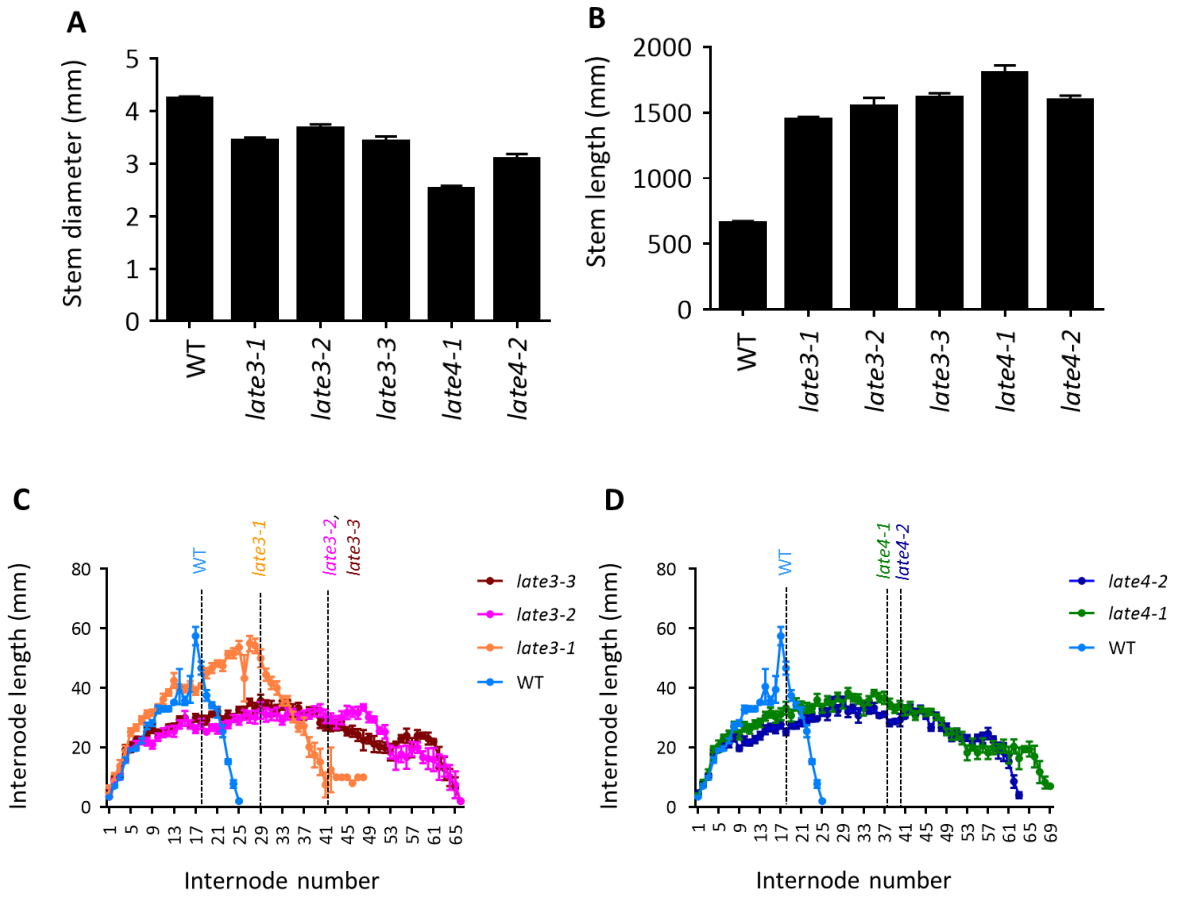


Figure 3.10. Variation in various traits of main stem in wild type (WT) NGB5839 and *late3* and *late4* mutants grown under LD. **(A)** Stem diameter, average of three independent measurement for internode 9, data represents mean ± SE for n = 6-10 plants. **(B)** Stem length refers to distance from node one to apex, data represents mean ± SE for n = 6 plants. **(C-D)** Internode length (IL), data represents mean ± SE for n = 6 plants. Internode one is defined as the length between first and second scale leaf, dashed lines represents initiation of flowering in respective genotype.

Internode length acts as an important index for plant response to exogenous stresses (Sibomana et al. 2013; McCall and Atherton 1995; Grimstad 1993). Various pea mutants showing variation in internode length have been previously categorized as GA synthesis, GA response and photomorphogenic mutants (Murfet 1993). Besides, characteristic reduction in internode length in the wild-type genotype during the time of flowering initiation is considered as a morphological indicator of the transition of shoot apical meristem (SAM) from vegetative meristem (V) to primary inflorescence meristem (I_1) (Sussmilch 2014). Due to these aforementioned reasons, the present study emphasized on studying in detail about the effect of mutation in *LATE3* and *LATE4* genes on internode length. The WT showed consistent increase in length between consecutive internodes with two peaks at internode 14 and internode 17 just prior to onset of first flower (Figure 3.10 C-D, WT). This was followed by a declining trend until plant growth was arrested due to senescence of apex. This result is in agreement with what was observed previously in similar studies in WT genotype (Ross et al. 1992a; Ross and Reid 1992; Sussmilch 2014). The pattern of ontogenetic variation in the weaker *late3* allele, i.e., *late3-1* was similar to WT with the first peak occurring at around same internode and the second peak appearing just prior to initiation of flowering at around 28 internode before falling off sharply in the following internodes until growth was stopped (Figure 3.10 C). Internode length between *late3-1* and WT was also more or less same until internode 18 when the WT flowered. In contrast, stronger alleles *late3-2* and *late3-3* had internodes of similar length with WT only at the initial stage of growth which was around internode six to seven (Figure 3.10 C). After that, both these mutants grew shorter internodes compared to WT until internode 21-22 by which time WT had already flowered. Besides, these mutants also did not exhibit a sharp upward trend in length from internode seven until emergence of flower at around internode 42, they rather formed more or less a plateau during this growth period and this was a clear difference in comparison to similar growth stage in the WT. Once flowering was initiated in these mutants, length in the following internodes did not show rapid decline which is typical in the WT genotype, instead a gradual descending trend was observed until the period of harvest, i.e. internode 67 (180 days after sowing) when the plants were still growing.

Growth pattern of internodes in terms of length in *late4-1* and *late4-2* was more or less similar to that observed in *late3-2* and *late3-3* (Figure 3.10 D). Nevertheless, the weaker *late4* allele,

i.e., *late4-1* had slightly longer internodes compared to stronger allele *late4-2* from internode 9 to approximately internode 40 when both these mutants flowered. Overall, the trend in alteration in internode length observed under LD conditions provide evidence that *LATE3* and *LATE4* are crucial for promoting internode length, and this may in part reflect an impaired transition from VM to I_1M stage.

3.4.7 LATE3 and LATE4 influence root development and overall biomass of the plant

Plant establishes itself in soil via the root as it uptakes water and various macro and micronutrients through this organ. The root architecture is also important for cross-talk with the microflora in the rhizosphere. In case of pea, it has even more importance because of potential symbiotic relationship with nodule bacteria which fixes atmospheric nitrogen into readily usable by the plant, i.e., ammonia (Tricot et al. 1997). Moreover, root is vital for adaptation to various abiotic stress conditions as well as defence against invading microbes. Since effect of *late3* and *late4* mutations were observed on various above ground vegetative organs, therefore the present study also examined various traits of the under-ground vegetative organ, i. e., root.

The WT had significantly larger primary root compared to the representative *late3* allele *late3-2* (Figure 3.11 A, $p < 0.05$). The *late4-1* mutants had also smaller primary root than WT, however the variation was not significant (Figure 3.11 B, $p = 0.2442$). In contrast, both *late3-2* and *late4-1* mutants produced lower root biomass in terms of root dry weight when compared to the WT which was not significant (Figure 3.11 B, $p = 0.057$). Nevertheless, these results suggest *LATE3* and *LATE4* seem to positively affect root length and biomass generation which in turn is related to nutrient and water acquisition from soil.

Nitrogen taken up by roots is delivered to photosynthetic source organs via phloem from which it is then translocated to sink organs such as pods and seeds (Zhang et al. 2015). Thus, biomass of the plant on early growth stage could provide valuable idea about yield potentiality on a later developmental stage. Since it was found that *late3* and *late4* mutants affect various yield related traits such as pod and seed development at late growth stage, so it was of interest to investigate whether these mutants also alter shoot biomass at early growth stage. As was observed, both *late3* and *late4* mutants exhibited significantly lower fresh weight of shoot in comparison to WT (Figure 3.11 C-D, $p < 0.05$). The dry weight of shoot

showed a similar effect as well. Based on these results as well as those from root biomass, it could be stated that *LATE3* and *LATE4* genes play role in overall growth of pea.

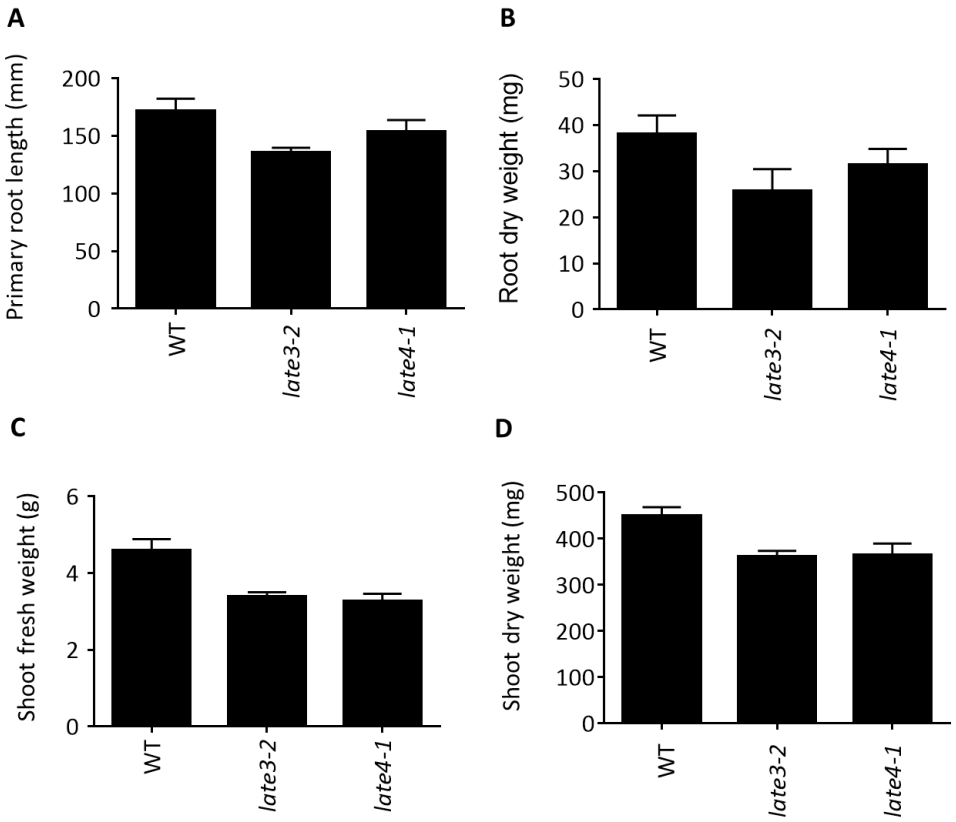


Figure 3.11. Characteristics of root and shoot biomass in wild type (WT) NGB5839 and *late3-2* and *late4-1* mutants grown under LD condition. **(A)** Primary root length measured 12 days after sowing, length covered distance from cotyledon to the tip of primary root, data represents mean ± SE for n = 8-11 plants. **(B)** Root dry weight, entire root mass was harvested 12 days after sowing and then dried for four days at 70 ° C before taking measurements, data represents mean ± SE for n = 8-11 plants. **(C)** Shoot fresh weight, this trait was measured immediately after harvest after 27 days of growth, data represents mean ± SE for n = 9-12 plants. **(D)** Shoot dry weight, samples were harvested after 27 days of growth and then dried for four days at 70 ° C before taking measurements, data represents mean ± SE for n = 9-12 plants.

3.5 Discussion

3.5.1 LATE3 and LATE4 are major regulators of flowering time and other relevant reproductive traits

As mentioned earlier, the flowering time mechanism was proposed to be conserved to a greater extent between the model *A. thaliana* and pea as a number of *A. thaliana* orthologous genes being involved in flowering have been identified in pea (Hecht et al. 2005). Delayed flowering phenotype was observed previously for other late flowering mutants such as *late1* (Hecht et al. 2007), *late2* (Ridge et al. 2016), *late5* (Sussmilch 2014) and *veg2* (Sussmilch 2014) all of which initiated flowering before node 30 and therefore the observed late flowering phenotype in *late3* and *late4* mutants (Figure 3.1 A-E) is the most severe form of this particular trait to date. These results suggest a strong positive regulatory role for *LATE3* and *LATE4* genes in pea flowering.

The present study revealed that many of the inflorescences and flowers generated in the l*ate3* and *late4* mutants had aberrant structure. Such abnormalities included in all the four whorls of the flower, i.e., sepals, petals, stamen and carpel of the dissected flower (Figure 3.4) and some of the non-dissected flowers also showed various types impaired development of different floral organs (Figure 3.3). The classical ABCDE model of floral organ identity in *A. thaliana* involves different types of MADS domain carrying genes which interact with each other in different ways in order to generate various organs of the flower (Theissen 2001). *A. thaliana* MADS domain genes such as *AP1*, *AP2,AP3*, *PI*, *AG, SEP1*, *SEP2*, *SEP3* and AGL11 are representatives of A-E groups (Theissen 2001). In pea, *PIM*/*PsAP1* act as a floral meristem identity gene (Taylor et al. 2002) whereas *DET* and *VEG1* are determinants of primary inflorescence meristem (I_1M) and secondary inflorescence meristem (I_2M) respectively (Berbel et al. 2012). Besides, *VEG2* was found to be expressed in all these tissues (Sussmilch et al. 2015). The results suggest that *LATE3* and *LATE4* genes are important for determining overall floral morphogenesis and they are likely to act in this regard through interaction with the aforementioned inflorescence and floral identity related genes.

Another important characteristic observed in the *late3* and *late4* mutants was their prolonged initial reproductive stage compared to WT (Figure 3.5 A). Role of various hormones such as

auxin, cytokinin, abscisic acid, ethlylene, jasmomic acid and salicyclic acid have been reported in the event of senescence in *A. thaliana* (Khan et al. 2014). Genome wide transcriptome analysis revealed that the network of miRNAs and *Agronaute1* enriched sRNAs which regulates senescence associated genes (SAGs) are quite conserved between *A. thaliana* and rice (Qin et al. 2016). The absence of flowers, pods and seeds along with continued vegetative growth in the mutants during a comparable time point with the WT would mean that while diversion of nutrients has taken place in the WT genotype, but this entire process is delayed in the mutants. It could be a fact that *LATE3* and *LATE4* might act upon the yet unidentified SAGs or modulate the regulation of the hormonal pathways in pea.

Rare biological events like inflorescence and floral reversion were noticed in the *late3* and *late4* mutants through this study (Figure 3.5 A-G). Previously, the process of inflorescence and floral reversion have been associated with inability to maintain floral commitment (Tooke et al. 2005; Adrian et al. 2009). The transcriptional network that regulates floral commitment in *A. thaliana* largely involves synergistic interaction between FM genes (*AP1*,*LFY*,*CAL)* and antagonism between *TLF1* and *AP1*/*LFY*/*CAL* in SAM and the *FT* mediated leaf born signal is unlikely to play vital role in this regard (Adrian et al. 2009). Under non-inductive SD conditions, heterozygous *lfy* and homozygous *ap1* mutants show floral reversion (Okamuro et al. 1996; Okamuro et al. 1997; Adrian et al. 2009). Besides, loss of function of *SOC1* and *FUL* is also known to cause inflorescence reversion (Melzer et al. 2008). In another species *Impatiens balsamina*, flowering commitment requires continuous supply of photoperiod mediated floral stimulus (Tooke et al. 1998; Tooke et al. 2005). Analysis of the *LFY*, *AP1*, *AG* and *TFL1* have shown the presence of general function of these genes, but absence of mutual regulation between them (Pouteau et al. 1997; Ordidge et al. 2005). In pea, it was also proposed that flowering is regulated by a balance between a promoter and inhibitor signal (Weller et al. 1997b). As *LATE3* and *LATE4* were found to be important for maintenance of floral commitment in the present study, therefore it can be hypothesized that mis-regulation of key pea flowering genes such as *FT*s, shoot meristem identity gene *LF*/*PsTFL1c* and inflorescence/floral meristem identity genes such as *VEG1*/*PsFULc, UNI*/*PsLFY* and *PIM*/*PsAP1* and also disturbance in their potential mutual interaction in *late3* and *late4* mutants might have caused inflorescence and floral reversion. Future research involving tissue samples from each of the stages of I-RN, IR-VN and L-RN (Figure 3.5) would unravel better understanding in

this regard. Potential epigenetic regulation of the aforementioned important pea flowering genes with relevance to *LATE3* and *LATE4* need to be investigated as well.

3.5.2 Various agronomically important traits of pods and seeds are controlled by LATE3 and LATE4 genes

In the present study, it was observed that only a small portion of the total flowers generated seed carrying pods in the *late3* and *late4* mutants (Figure 3.6 A-B). Majority of these looked like a normal WT flower morphologically while a certain portion of them had defective floral organs as discussed above. In plants, the process of viable gamete formation through micro- and megasporogenesis is important for successful fertilization and these processes are controlled by various genes involved in pre-meiotic to post-meiotic mitosis (Pagliarini 2000). The MADS box TF known as SPL/NZZ is crucial for male and female gametogenesis in *A. thaliana* (Wilson and Yang 2004). Unlike animals, the sperm cells in angiosperms are immobile and they are transported to the egg apparatus by the formation of pollen tube. Once the two sperm cells are released within the female gametophyte/embryo sac, then double fertilization takes place through their interaction and fusion with the egg and central cell which generates integral seed components embryo and endosperm respectively (Bleckmann et al. 2014; Dresselhaus et al. 2016). In *A. thaliana*, transmembrane receptor like kinases such as FER are linked to synergid cells while Ca^{2+} transporting proteins such as ACA9, GLRs are involved in pollen tube growth (Bleckmann et al. 2014; Dresselhaus et al. 2016). Similarly, secreted small cysteine rich peptides such as EC1 are associated with egg cells whereas single-pass transmembrane proteins such as GEX, HAP2 are specific to sperm cells. All these components along with some other factors interact with each other for successful completion of the process of double fertilization. After fertilization, the early stages of embryogenesis involves several divisions of the zygote in *A. thaliana* to generate an octant where leucine zipper TF class IV ATML1 and its closely related protein PDF2 and leucine rich repeat receptor like kinases RPK1 play vital role (Abe et al. 2003; Nodine et al. 2007; Takada and Jurgens 2007; ten Hove et al. 2015). Based on the results of the current study, it can be speculated that the impairment in the process of gametogenesis/fertilization/early embryogenesis may have occurred in both *late3* and *late4* mutants because of which much of the flowers having WT outlook senesced without proceeding to the step of seed formation. *LATE3* and *LATE4* may act in the regulatory process that drive various steps of gametogenesis, fertilization and early

embryogenesis in pea. Follow-up studies involving detailed meiotic analysis (Pagliarini 2000), pollen viability test (Shivanna and Rangaswamy 1992; Pagliarini 2000), live cell imaging of double fertilization using male and female gamete specific protein markers (Berger et al. 2008), two photon microscopic observation of sperm cell movement (Hamamura et al. 2012) and computational analysis of cellular patterning of embryo development in 3D (Yoshida et al. 2014) in WT and mutant genotypes is likely to reveal better knowledge about this process.

In the present study, *late3* and *late4* mutants showed modulation with WT for various important traits such as seed number/pod, size and dry weight as well as pod shape and size (Figure 3.6, 3.7) which gives hint that the underlying regulatory processes controlling these traits might be affected in the mutants. All the tissue layers that are needed for developing a mature fruit are formed within the ovary before fertilization where the ovules and ovary wall are transformed into seed and fruit cover respectively (Dinneny and Yanofsky 2005; Robles and Pelaz 2005). In most species, the process of fruit development occur simultaneously along with seed development which is largely dependent on acquiring relevant signals from the developing seeds (Sotelo-Silveira et al. 2014) and information on this process in pea is already available (Ozga and Reinecke 2003). The chlorinated auxin namely 4-chloroindole-3-acetic acid (4-Cl-IAA) is found naturally in pea seeds and two amino transferases namely TRYPTOPHAN AMINO TRANSFERASE RELATED1 (PsTAR1) and PsTAR2 have crucial role in biosynthesis of 4-Cl-IAA (Reinecke 1999; Tivendale et al. 2012). 4-Cl-IAA is transported from the developing seed to the ovary (pericarp) which modulate GA and ethylene biosynthesis pathway and thereby regulate fruit growth and development (Ozga and Reinecke 2003; Jayasinghege et al. 2017). In *A. thaliana*, the floral meristem identity gene *FUL* is expressed from early stages of flower development to late fruit development as it play role in fruit patterning (Robles and Pelaz 2005; Ferrandiz et al. 2000). Mutation in *AtFUL* results in reduction of silique size. In pea, three copies of the *FUL* genes were found and among them *VEG1*/*PsFULc* gene also act as a floral meristem identity gene (Berbel et al. 2012), but its role in relation to seed and fruit development has not been investigated so far. Based on these information, it can be speculated that mutation in *LATE3* and *LATE4* genes might affect function of the pea homologs of the aforementioned *A. thaliana* genes important for seed number, size and development. In addition, the 4-Cl-IAA and potentially also *PsFULc* mediated fruit development process in pea may also be affected. Future studies such as expression

experiment using samples from different stages of seed and fruit development is likely to unravel better understanding about these crucial yield related processes in pea.

3.5.3 LATE3 and LATE4 affect diverse vegetative traits

For any given plant species specially for crop plants such as pea, various vegetative traits of leaf, shoot and root during the different growth stage are likely to play important role in ultimately determining the status of agronomically important traits such as yield. In this project, studies were carried out at various time points of development for different vegetative traits in *late3* and *late4* mutants which revealed marked difference relative to the wild type NGB5839.

Shoot branching is a complex trait and to date a very complicated, however interlinked network involving auxin, cytokinin, strigolactone, gibberellic acid, abscisic acid, sugar biosynthesis have been placed forward as the model of shoot branching in the model plant *A. thaliana* (Rameau et al. 2015). Studies of mutants involving molecular physiology techniques in *A. thaliana* as well as in other species such as pea led to the identification of a conserved mechanism whereby a long distance mobile signal considered to be derived from strigolactone which is regulated by two genes known as *Carotenoid cleavage dioxygenase 7* (*CCD7*) and *CCD8* (Sorefan et al. 2003; Booker et al. 2004; Johnson et al. 2006; Simons et al. 2007; Snowden et al. 2005) play a major role to inhibit shoot branching. Thus, both of these genes are crucial in shoot branching further downstream of auxin. In pea, homologs of *CCD7* and *CCD8* are known as *RMS5* and *RMS1* respectively (Rochange 2010; Morris et al. 2001). Specific role of strigolactones in this process was proposed where it downregulates the function of a pea shoot branching promoting gene *BRC1* (Braun et al. 2012). Trehalose 6 phosphate (T6P) which serves as a marker for sugar availability within the plant has been found to promote axillary bud outgrowth in pea (Fichtner et al. 2017). During the current study, the *late3* and *late4* mutants (Figure 3.8) demonstrated massive aerial branching and basal branching (only in *late4* mutants) unlike any other previously characterized late flowering mutants (section 3.4.5). It is likely that alteration of sugar allocation in the lateral bud or hormonal biosynthetic process which is affected by the mutation in *LATE3* and *LATE4* genes might be playing a role in this regard.

In the present study, modulation in various leaf related traits were observed in the *late3-2* and *late4-1* mutants (Figure 3.9 B-D). In *A. thaliana*, leaf size is determined by genes involved in various processes such as size of primordium, rate of cell division and expansion and meristemoid division (Gonzalez et al. 2012). Characterization of the *uni* mutant in pea had revealed that the underlying gene which is an orthologoue of *A. thaliana LFY* is involved in leaf morphogenesis (Hofer et al. 1997). Another study identified the relative effect of two genes *Af* and *Tl* on pea leaf development (DeMason and Villani 2001). Moreover, role of plant growth hormones such as auxins and gibberellin has also been attributed during various stages of leaf development (Demason and Chawla 2006). Based on these information, it can be hypothesized that *LATE3* and *LATE4* genes might act in any of the potentially conserved or pea specific regulatory pathways that regulate size of the various components of the compound leaf.

In plants, differential activity of the apical and lateral meristems causes flexibility in stem growth leading to variability in shoot architecture which is most notable example of adaptation (Sanchez et al. 2012). Secondary growth in plants generates secondary tissue such as secondary xylem from vascular cambium. This in turn causes an increase in stem diameter and circumference (Oh et al. 2003). The present study revealed that *late3* and *late4* mutants had smaller stem diameter compared to WT (Figure 3.10 A). This trait is important for pea in order to prevent canopy collapse. Various hormones such as auxin, cytokinin, ethylene, gibberellins and jasmonate are known to promote cambium initiation and activity as well as secondary growth (Bjorklund et al. 2007; Hejatko et al. 2009; Love et al. 2009; Sehr et al. 2010; Agusti et al. 2011). Research involving *A. thaliana* have revealed that the two key flowering genes *SOC1* and *FUL* act as repressors of cambium formation (Lens et al. 2012). It can therefore be speculated that *LATE3* and *LATE4* promote stem thickness acting through any of the aforementioned regulatory pathways and thus they are probably crucial for transportation of nutrients via stem.

Ontogenetic variation of internode length has been identified as a key trait for determination of plant transformation from vegetative to reproductive stage (VM/I_1M) and role of gibberellin is quite known in this regard. Previously, measurement of the concentration of GA1 at various internodes revealed higher concentration of the hormone in internodes until NFI and a marked decrease in the following internodes (Ross and Reid 1992; Ross et al. 1992b).

Auxins such as Indole-3-acetic acid (IAA) shifts from the growing apex to internode to mediate its elongation via the acceleration of GA1 biosynthesis and accumulation as well as masking deactivation of GA1 (Ross et al. 2003; Ross et al. 2000; Weston et al. 2009). Pea late flowering mutants having shorter internodes such as *veg1* and *gigas* were used for quantification of GA1 and IAA both of the hormones were present in low concentration in the mutants (Beveridge et al. 2001). *VEG1* and *GIGAS* have been identified as being involved in VM/I_1M transformation and acting as a mobile floral promoting signal respectively (Hecht et al. 2011; Berbel et al. 2012). In another study, *veg-2* mutant also showed shorter internodes compared to WT where the underlying gene was identified as pea homologue of *AtFD* (Sussmilch et al. 2015). It was proposed that a link between the hormonal pathway and the V/I_1 transition pathway could be potential reason for the observed phenotype where the plant commits to strengthen reproductive success by modulating hormonal levels. Previously, role of *LF* in pea was established which act as premier regulatory gene in determining the length of vegetative stage and it acts as an inhibitor of flowering initiation (Foucher et al. 2003). During the current study, the *late3* and *late4* mutants did not just show variation in internode length, rather they also exhibited prolonged vegetative growth and longer overall shoot length in compared to the WT (Figure 3.10 B-D). Besides, they did not exhibit (apart from *late3-1*) the characteristic reduction in internode length at onset of flowering which occurs in WT and thereby indicates that the SAM has underwent VM to I_1M transformation. Therefore, it is likely that *LATE3* and *LATE4* act through any of the aforementioned hormonal pathways in regulating internode length and/or any of the already known flowering genes in promoting the transition from vegetative to reproductive stage.

In the present study, *late3-2* mutant, but not *late4-1* showed shorter primary root providing hints for a key role of *LATE3* gene in promoting root elongation (Figure 3.11 A). On the other hand, both *LATE3* and *LATE4* genes are likely to be involved in positively formulating overall structure of root biomass as the dry weight of root was lower in both the mutants compared to WT (Figure 3.11 B). Various plant growth hormones such as auxin, cytokinin, brassinoids, gibberellin, abscisic acid and strigolactones are known to play role at various steps of root development (Bishopp et al. 2011; Yu et al. 2011b; Brewer et al. 2013; Shani et al. 2013; Pacifici et al. 2015; McAdam et al. 2016). In pea, it was reported previously that DELLA proteins LA and CRY aid in expression of GA synthesis genes and supress root elongation

(Weston et al. 2008). A possible role of *LATE3* and *LATE4* genes in any of these broader regulatory network which is responsible for generation and elongation of primary and lateral roots is most likely to be the reason of observed effect in the mutants.

The current study showed a significantly lower fresh and dry weight of shoot in the *late3* and *late4* mutants relative to WT during an early growth stage (Figure 3.11 C-D) indicating likely effect of such traits in determining yield at a later developmental stage. The inorganic nitrogen which is taken up by the roots from soil is converted into amino acids within the root cells in pea (Pate 1980; Oaks 1992). These amino acids are then transported to leaves via the transpiration stream in phloem. Quantity of flowers, pods, seeds and storage proteins within the seeds are vastly dependent on the transport of nitrogen from leaves to the sink tissues (Crawford 1995; Forde 2002; Sanders et al. 2009; Schmidt et al. 2007; Tan et al. 2010). Besides, role of amino acid transporter namely acid permease AAP2 in improving biomass by increasing xylem to phloem transfer was established in *A. thaliana* in a previous study (Zhang et al. 2010). Based on this information, *LATE3* and *LATE4* genes could promote shoot biomass accumulation by interacting with these genes. It may also be a cumulative effect of regulating the development of different organs, i.e., root, stem and leaf at the early growth stage by *LATE3* and *LATE4*.

3.5.4 Concluding remark

The phenotypic results of this chapter revealed that *LATE3* and *LATE4* genes regulate a range of reproductive, vegetative and yield related important traits during various growth stages of the plant providing hint about their potential role as global regulatory genes. Nevertheless, the various observed phenotype of the mutants may also be a pleiotropic effect of the late flowering phenotype. Observation of similarities in new phenotypes and confirmation of same phenomenon for previously studied phenotypes (Weller et al, 2007) between *late3* and *late4* mutants conducted during the current project assisted in solidifying hypothesis that the underlying genes probably work through mutual interaction in the same biological pathway.

Chapter 4: Genetic mapping of *LATE3* and *LATE4* loci

4.1 Introduction

4.1.1 Genetic markers and genetic maps

Genetic markers could be termed as specific DNA sequences with known physical location within the chromosome of a species which represent genetic variation between different genotypes (Collard et al. 2005; Benavides and Guénet 2012; Nadeem et al. 2018). Thus, genetic markers that are tightly linked/located within close proximity of the gene of interest act as signs or tags and assist scientists to get near to the actual gene mediating the phenotype under investigation (Collard et al. 2005). DNA/molecular markers are the most widely used genetic markers. The characteristic feature of good quality DNA marker is codominant inheritance (ability to distinguish between homozygotes and heterozygotes), high reproducibility and neutral behaviour (Collard et al. 2005).

In order to develop genetic map for a locus controlling a trait of interest, it is necessary to generate a segregating mapping population by crossing two parental genotypes that vary from each other for that particular trait (Collard et al. 2005). For a preliminary genetic map, usually 50-250 F_2 individuals are needed, but even larger population is required for creating high-resolution map (Mohan et al. 1997; Collard et al. 2005). The principle of genetic mapping underlies the event of segregation between markers through chromosomal recombination/crossing-over during meiosis which can be analysed in the resultant progenies (Collard et al. 2005). Genes and markers that are located nearby are likely to be linked and transferred more commonly to progeny from parent than those which are placed further apart. DNA markers showing polymorphism between the parents initially are then needed to be screened (a process called genotyping) in the entire mapping population and F_1 heterozygote (if possible) as well in order to determine the genotype of the marker loci in those individuals (Collard et al. 2005). Each individual of the population is given a code based on relevant genotype for each marker (e.g. Parent 1, Parent 2 or F_1 heterozygote) and such data is used for estimating recombination frequency between markers. Thus, a linkage map which exhibit the position and relative genetic distance between markers in the chromosome

is generated. Higher the recombination frequency (known as map units centi Morgan or cM) between two markers, greater is the genetic distance between them and vice-versa. As it is not feasible to determine linkage between large numbers of markers manually, therefore various computer programs are regularly used for this purpose. In this case, estimation of linkage between markers is done by calculating odds ratio (linkage vs. no linkage) expressed as logarithm of odd value (LOD) (Risch 1992). Commonly used threshold for generating linkage map is LOD value of >3 (Collard et al. 2005). A LOD value of 3 denotes that the ratio of likely linkage : non-linkage between two markers is 1000 : 1 (null hypothesis). Most widely used softwares for developing linkage maps include MapManager QTX (Manly et al. 2001), MapMaker/EXP (Lincoln et al. 1993) and JoinMap (Stam 1993).

4.1.2 Status of molecular tools for genomics studies in pea and preliminary mapping of LATE3 and LATE4 loci

Genetics studies started in pea as early as the 1790s (Ellis 2009) which then led to the discovery of the two laws of inheritance by Mendel (Mendel 1865) and thereby laying the foundation of the principles of genetics. Despite being used as a plant species of early genetic studies and a very important crop for food security, pea has received much less attention than many other crop species (Tayeh et al. 2015b). Hence, availability of applied genomic and bioinformatics tools in pea is still limited compared to other major crop group such as cereals (Bhattacharyya et al. 1990; Hofer et al. 2009). The presence of higher degree of transposable elements as well as the larger genome size of pea probably the main reason in the hindrance of such developments (Neumann et al. 2001; Macas et al. 2007). However, genome sequencing of pea is underway by an international consortium, and it is not yet accessible publicly (Madoui MA 2015). In contrast, genomics resources in the form of genome sequence have been developed in other model legume species such as *Medicago truncatula* largely due to the convenience of dealing with its small genome size of about 500 Mbp (Sato et al. 2007). Knowledge gained from this species for various complicated developmental processes are routinely being translated in order to gain insight about similar process in higher order legume species such as pea (Young and Udvardi 2009; Sanders et al. 2011).

As mentioned in chapter 3, forward genetics based approach have been exploited to date for identification of various developmental genes. To this end, various functional maps consisted of genes of known function (Aubert et al. 2006; Bordat et al. 2011) and high resolution

consensus genetic maps (Sindhu et al. 2014; Tayeh et al. 2015a) for all of the pea linkage groups having information on close syntenty with different *M. truncatula* chromosomes have been developed which serve as the basis of recent genetic mapping studies in pea. Thus, comparative genomics studies via exploitation of synteny between pea and the closely related species *M. truncatula* has been a key technique in developing genetic maps and candidate gene identificaton for various types of mutants in pea such as those involved in flowering, branching, light and ethylene signalling (Weller et al. 2009; Liew et al. 2014; Sussmilch et al. 2015; Weller et al. 2015; Ridge et al. 2016; Rubenach et al. 2017; Ligerot et al. 2017).

Even though genomic and expressed sequenced tag (EST) based short sequence repeat (SSR) or microsatellite markers were generated across all the seven linkage groups of pea through various studies (Ford et al. 2002; Loridon et al. 2005; Sun et al. 2014; Burstin et al. 2001; Gong et al. 2010; Kaur et al. 2012; Mishra et al. 2012), the single nucleotide polymorphism (SNP) markers proved to be more useful. The reason behind this is their high abundance, biallelic and codominant nature, ease in automatic genotyping and amenability in cross species studies and several SNP based high resolution, consensus map have been developed so far (Leonforte et al. 2013; Duarte et al. 2014; Sindhu et al. 2014; Tayeh et al. 2015a). Introns are of key focus for the generation of such SNP markers since the likelihood of getting polymorphism was found to be as much as three times higher than the exonic region in studies involving human (1 SNP/1000bp) and soybean (1 SNP/200-300bp) (Wang et al. 1998; Zhu et al. 2003). Likewise, 1 SNP in every 20 bp was reported for the intronic region in pea in another study (Jing et al. 2007).

Narrow cross using mutant loci in the background of NGB5839 and another cultivated pea line cv. Terese have been successfully used in previous studies in order to refine map position of the relevant flowering loci which assisted in cloning the underlying gene (Hecht et al. 2007; Sussmilch 2014). An F_2 population of 100-120 progenies is considered to be sufficient for initial localization of the gene of interest using five to seven markers along each of the pea linkage groups which potentially generates genetic map having distance of around 30-35 cM between each markers (Weller et al. 2013).

For the purpose of mapping *LATE3* locus, the pea cultivated line cv. Terese was used to generate segregating F_2 mapping population through cross with *late3-1* mutants (Weller et.

al, unpublished data). Likewise, *late4-1* x cv. Terese and *late4-2* x cv. Terese were used for developing F_2 progenies in order to map *LATE4* locus. Initial results of linkage studies placed *LATE3* locus at the middle of *Pisum sativum* linkage group III (*PsLGIII*) and *LATE4* locus at the bottom of *PsLGV* (Hecht et. al., unpublished data) which are syntenic to *M. truncatula* chromosome 3 and 7 respectively (Kaló et al. 2004).

4.2 Chapter aims

In general, the goal of this chapter was to investigate the *LATE3* and *LATE4* loci further in detail by exploiting syntenic relationship between pea and *M. truncatula*. As part of that, initial aim was to improve the genetic map by re-analysing the phenotype and marker data of the existing mapping populations. The next goal was to refine map position of both the loci through generation of supplemental F_2 mapping population as well as adding more markers.

4.3 Materials and methods

Various materials and methods used specifically for this chapter are mentioned here. The general materials and methods exploited in this chapter are given in chapter 2. Details of primers are provided in appendix 1.

4.3.1 Generation of new mapping population

Plants for this chapter were grown at 20° C (16 hour photoperiod) in the top phytotron of the School of Natural Sciences, University of Tasmania.

As part of investigation in this chapter, new F_2 mapping population consisting of 108 individuals were generated from cv. Terese x *late3-1* under LD condition (18 hour photoperiod) in the top phytotron of the School of Natural Sciences, University of Tasmania. Similarly, 78 F_2 plants were developed through *late4-1* x cv. Terese. Seeds for sowing were obtained from J. Weller and J. Vander Schoor. DNA from both the aforemetioned populations were extracted according to the prototocol mentioned in chapter 2 which were then diluted to a concentration of 50ng/µl for subsequent usage after measuring their concentration in nanordrop 8000 spectrophotometer (Thermosfisher Scientific, Wilmington, DE).

Data on node of flowering initiation (NFI) was collected for all of these new F_2 plants. The homozygous late flowering segregants could be clearly distinguished from the early flowering homozygous/heterozygous segregants based on NFI. In order to differentiate between early flowering homozygous and heterozygous dominant segregants for the classical marker, selected F_2 plants from both Terese x *late3-1* and *late4-1* x Terese populations were further grown in F_3 generation. Genotype of the respective F_2 individuals were considered as homozygous cv. Terese for phenotypic marker *LATE3*/*LATE4* if early flowering segregant appeared in an F_3 family size of $n > 8$ or heterozygous cv. Terese for phenotypic marker *LATE3*/*LATE4* if no early flowering segregant was obtained in the same sample size. If the sample size was $n < 8$ having no early segregants, then the F_2 plants were labelled as undetermined. Thus, four categories of codes were used for linkage analysis using Joinmap 4 (Kyazma B.V., Wageningen, Netherlands). Using this segregation data from F_2 population, linkage map was obtained as an estimation of genetic distance between the markers.

4.4 Results

4.4.1 Refinement of map position for LATE3 locus

Prior to the commencement of this study, *LATE3* locus was mapped using an F_2 mapping population consisted of 192 individuals generated via a cross between cv. Terese x *late3-1.* The resultant map placed *LATE3* locus within a 38.9 cM region between markers *AAP2* and *RKP1* near to the base of *PsLGIII* (Hecht et al, unpublished) (Figure 4.1 A, Table 4.1). This initial study also revealed almost a Mendelian segregation ratio of 3:1 for wild type and mutant flowering phenotype respectively in the F_2 population (Weller et. al, unpublished). DNA from this population were available, but several issues with this older study limited its usefulness which are as follows: (i) large proportion of missing data in marker genotyping, (ii) ambiguous scoring of phenotypes in some cases, (iii) low quantity and quality of stored DNA.

In order to overcome the shortcomings stated above, the current study initially focused on correcting marker data where possible by repeating genotyping of some of the existing markers in the already available F_2 individuals. Any phenotypic data that appeared erroneous were also discarded, e.g., node of flowering initiation (NFI) was not recorded while phenotyped as early flowering or genotyping was done without recording the phenotype data. Finally, the original population was supplemented by additional F_2 population of 108 plants which was available to be grown and genotyped in the present study. Individuals of this population showed clear segregation of the mutation based on delayed flowering and other characteristic traits as described in chapter 3 (Figure 4.2). Out of the total number of 300 individuals, 45 were not of further use after re-analysis of the original data and lack of germination in the additional population. This round of marker data analysis resulted in improvement of the original map due to alteration in the order of markers (Figure 4.1 B). To further refine position, new gene based markers were also developed by exploiting synteny between pea and *M. truncatula.* Since Hecht et al. (unpublished) identified map position of *LATE3* locus within a broader region between *NIP2* and *RKP1* markers in *PsLGIII*, therefore various *M. truncatula* genes possessing introns within the syntenic region of chromosome 3 (Kaló et al. 2004) were chosen for designing primers for further marker development in order to get near to the *LATE3* locus (Table 4.1, Table A1.1). The markers were developed at the exon-intron borders on pea transcripts after being aligned with orthologous *M. truncatula*

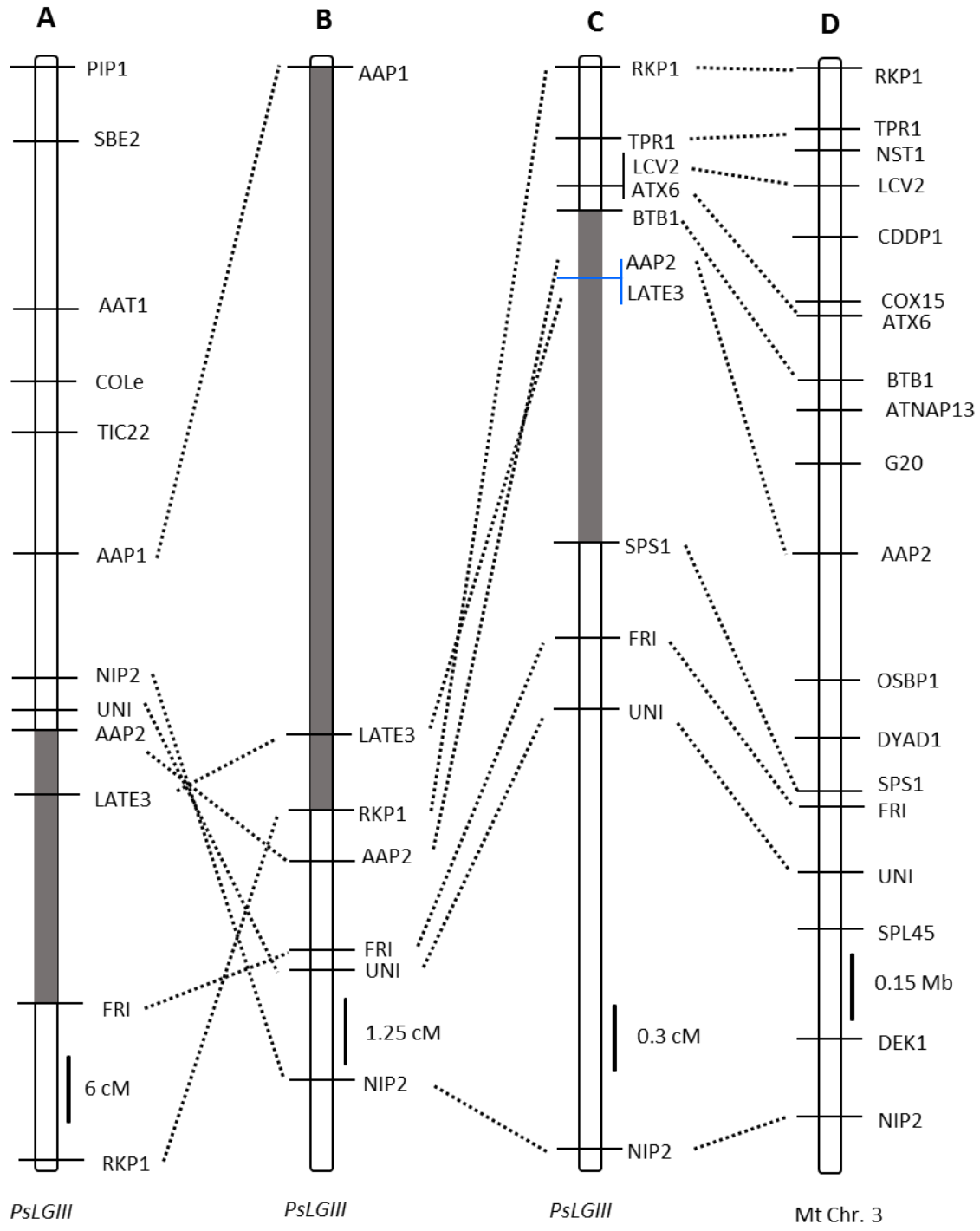


Figure 4.1. Various steps in the development of genetic map for *LATE3* locus using F_2 mapping population of cv. Terese X *late3-1*. **(A)** Previous mapping of *LATE3* in *P. sativum* linkage group III by Hecht et al. (unpublished) using 192 F_2 individuals. **(B)** Improvement of the original map after reanalysis of original data and supplementation with data from an additional F_2 population making a total of 255 individuals. **(C)** Fine mapping of the *LATE3* locus by adding new markers. **(D)** Comparative physical map of *M. truncatula* in chromosome 3 showing position of all the genes targeted for marker development. Shaded region represent defined interval showing location of *LATE3* locus between flanking markers.

transcripts. Through PCR amplification in NGB5839 and Terese, if polymorphism was obtained between these genotypes for the sequenced PCR product, then attempts were made to generate a molecular marker around the site of polymorphism and then scored in the entire F_2 population. This approach resulted in the development of five new markers out of 21 genes being used in an attempt to develop such markers (Figure 4.1 C, Table 4.1, A1.1, A1.3). The dual strategies of repeating genotyping previous markers and developing new markers in the larger F_2 population refined the map position of *LATE3* locus in a region of 1.5 cM between new markers *BTB1* and *SPS1* at the bottom of *PsLGIII* (Figure 4.1 C). This region was syntenic to 0.9 Mb in chromosome 3 of *M. truncatula* (Figure 4.1 D, Table 4.1).

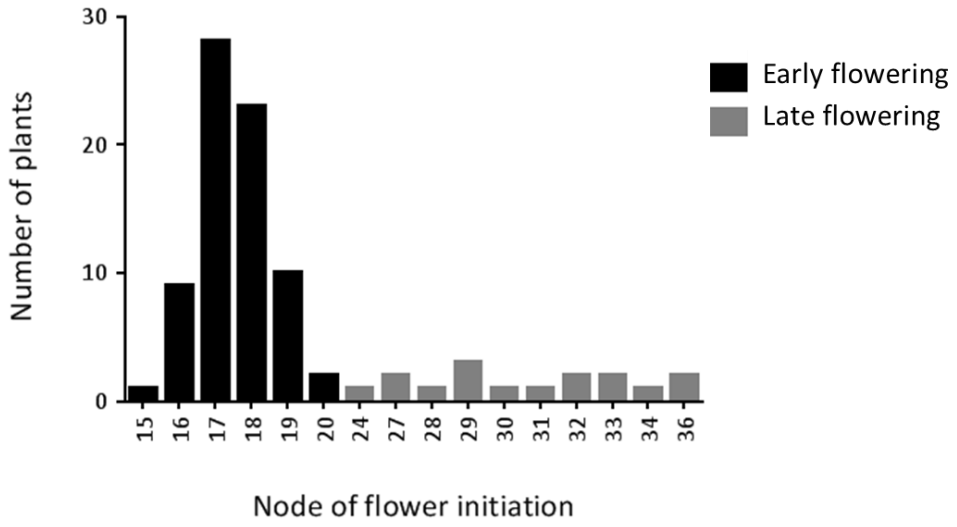
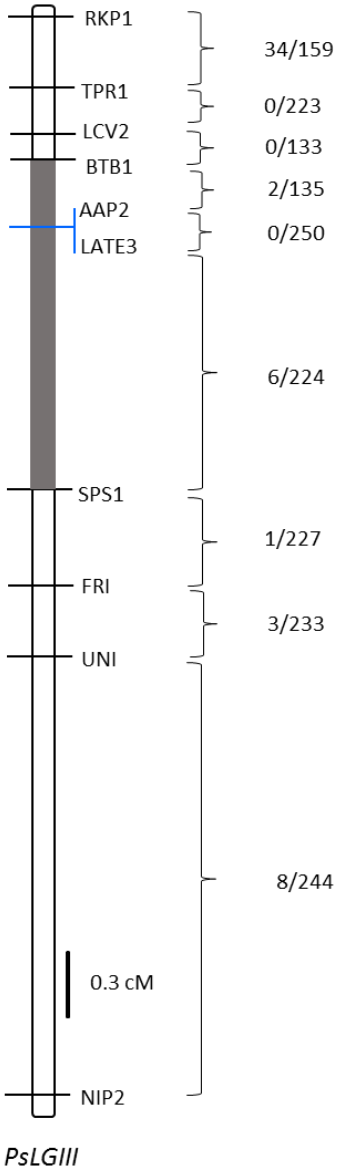


Figure 4.2 Segregation for node of flower initiation among the F_2 plants from cv. Terese X *late3-1* grown under LD condition during the period (2015) of this study.

Careful analysis of recombinants between neighbouring markers led to the identification of six recombinants showing segregation between *LATE3* and *SPS1* whereas there were two recombinants that showed segregation between *BTB1* and *LATE3* (Figure 4.3 A). Analysis of these eight recombinants also revealed that the initially mapped marker *AAP2* co-segregates with the *LATE3* locus (Figure 4.3 A-B). For the purpose of further refinement of the data, F_3 progeny from relevant recombinants, e.g., recombinant 2, 7 (Figure 4.3 B) having WT phenotype were grown (if seeds were available) in order to determine whether classical (*LATE3* locus) marker genotype for those individuals would be homozygous or heterozygous

A



B

	Recombinant 1	Recombinant 2	Recombinant 3	Recombinant 4	Recombinant 5	Recombinant 6	Recombinant 7	Recombinant 8
AAP1	B	H	H	A	-	B	B	H
RKP1	H	H	H	A	H	H	B	B
TPR1	H	H	H	A	H	H	B	B
ATX6	H	H	-	A	-	-	B	B
LCV2	H	H	H	A	H	H	B	B
BTB1	H	H	H	A	H	H	B	B
AAP2	A	H	H	A	H	H	B	H
LATE3	A	H	C	A	C	C	B	C
SPS1	A	B	A	H	A	A	H	H
FRI	A	B	A	H	A	A	H	H
UNI	A	B	A	H	A	A	H	H
NIP2	A	B	A	H	A	A	H	H

Figure 4.3. Analysis of recombinants that showed segregation between neighbouring markers using genotype data from F_2 mapping population of cv. Terese X *late3-1*. **(A)** Number of recombinants out of total genotyped individuals between neighbouring markers. **(B)** Genotype of important recombinants used to determine potential map position of *LATE3* locus. Genotype code used according to the instruction of JoinMap 4 software; A = Homozygous *late3-1,* B = Homozygous cv. Terese, H = Heterozygous, C = Homozygous cv. Terese / Heterozygous. Marker name colored with red were genotyped only in the important recombinants.

(Figure 4.3 B). Sowing seeds from recombinant 3, 5 and 6 (Figure 4.3 B) in F_3 was not essential as markers from *NIP2* to *SPS1* showed linkage in these individuals with all possessing homozygous mutant genotype. These recombinants showed early flowering phenotype and

Table 4.1. Genes attempted for developing markers to map *LATE3* locus in cv. Terese x *late3-1* F2 population in this study. *M. truncatula* Mt4.0v1 database and Pea RNA seq gene atlas database were used for collecting accession numbers. Information on primers and PCR carried out to obtain polymorphism between parental genotypes are given in table A1.1 whereas primers for markers are mentioned in table A1.3.

Locus name	*P. sativum* gene accession number	*M. truncatula* gene accession number	Physical position on *M. truncatula* chromosome 3 (Mb)	Marker type	Marker reference
AAP1	PsCam048093	Medtr3g088455	40.2	CAPS	Aubert et al., 2006
RKP1	PsCam026800	Medtr3g094770	43.26	HRM	Hecht et al., unpublished
TPR1	PsCam038860	Medtr3g095070	43.39	HRM	This study
WRKY13	PsCam039266	Medtr3g095080	43.41	No marker	-
NST1	PsCam023354	Medtr3g095130	43.45	No marker	-
LCV2	PsCam021197	Medtr3g095210	43.51	HRM	This study
CDDP1	PsCam037008	Medtr3g095470	43.62	No marker	-
COX15	PsCam043034	Medtr3g095790	43.76	No marker	-
ATX6	PsCam038017	Medtr3g095840	43.8	dCAPS	This study
BTB1	PsCam000088	Medtr3g096160	43.94	HRM	This study
ATNAP13	PsCam034645	Medtr3g096300	44	No marker	-
G20	PsCam046205	Medtr3g096500	44.11	No marker	-
AAP2	PsCam049915	Medtr3g096830	44.32	Size	Aubert et al., 2006
CR88	PsCam021158	Medtr3g096860	44.32	No marker	-
CDK8	PsCam048317	Medtr3g096960	44.39	HRM	This study
OSBP1	PsCam059306	Medtr3g097270	44.6	No marker	-
DYAD1	PsCam049636	Medtr3g097570	44.73	No marker	-
PPT1	PsCam035950	Medtr3g098060	44.74	No marker	-
SPS1	PsCam045416	Medtr3g098260	44.84	HRM	This study
FRI	PsCam036412	Medtr3g098290	44.86	Size	Hecht et al., unpublished
UNI	PsCam056640	Medtr3g098560	45	CAPS	Hecht et al., unpublished
SCPL45	PsCam034080	Medtr3g498755	45.12	No marker	-
DEK1	PsCam049636	Medtr3g098990	45.36	No marker	-
NIP2	PsCam013125	Medtr3g099380	45.55	CAPS	Aubert et al., 2006

the genotype of the classical marker is likely to be heterozygous as all the markers from *AAP1* to *AAP2* showed linkage having heterozygous genotype. In order to have homozygous WT genotype, a double recombination event is needed to occur within the short genetic distance of 1.5 cM between *BTB1* and *SPS1* which seemed very unlikely. Recombinant 1 (Figure 4.3 B)

showing late flowering phenotype have homozygous mutant genotype for the classical marker and it was in linkage with markers *AAP2* to *NIP2*, all of which also had homozygous mutant genotype. Overall, results from the new mapping data delivered strong evidence of the position of *LATE3* locus within a very small region in *PsLGIII*. Candidate gene selection from the syntenic genes in *M. truncatula* chromosome 3 (shaded region in Figure 4.1 C, 4.3 A) is discussed in the next chapter.

4.4.2 Refinement of map position for LATE4 locus

Similar to *LATE3* locus, Hecht et al. (unpublished) also initiated the mapping for *LATE4* locus in an F_2 population of 89 individuals generated from a cross between *late4-2* and cv. Terese. This preliminary map placed *LATE4* locus at the bottom of *PsLGV* with a genetic distance of 20.9 cM from the nearest marker *WRI11* on top, and there was no flanking marker below *LATE4* (Figure 4.4 A). This portion of *PsLGV* is syntenic to *M. truncatula* chromosome 7 (Kaló et al. 2004). Phenotypes of the F_2 population also showed nearly a Medelian segregation ratio of 3:1 for early flowering wild type and late flowering mutant segregants respectively (Weller et al., unpublished). However, this population also suffered from the same problems described earlier for the *late3-1* population and could not be used for further analysis. Instead, a second F_2 mapping population comprised of 111 individuals generated through cross between *late4-1* and cv. Terese was used. This population was previously grown and phenotypic data as well as stored DNA were of good quality (Weller et al., unpublished). Mapping some of the existing markers in this new F_2 population confirmed the position of the *LATE4* locus at the bottom of *PsLGV* (Figure 4.4 B).

In addition to mapping existing markers in the available *late4-1* x Terese population, process of designing new gene based markers by exploiting synteny with chromosome 7 of *M. truncatula* was also initiated as mentioned for *LATE3* in section 4.4.1. Based on the position of markers observed in the initial map generated by Hecht et al. (unpublished, Figure 4.4 A), all the *M. truncutula* genes located in a physical position from the marker *WRI11* towards the bottom of chromosome 7 were extracted. NGB5839 and cv. Terese DNA were used for PCR and subsequent quest for detection of polymorphism (Table 4.2, Table A1.2). Design of these PCR primers was similar in approach to that stated earlier for mapping of *LATE3* locus. Altogether, 17 new markers were developed in total using 54 primer combinations on different genes (Table A1.2, A1.3).

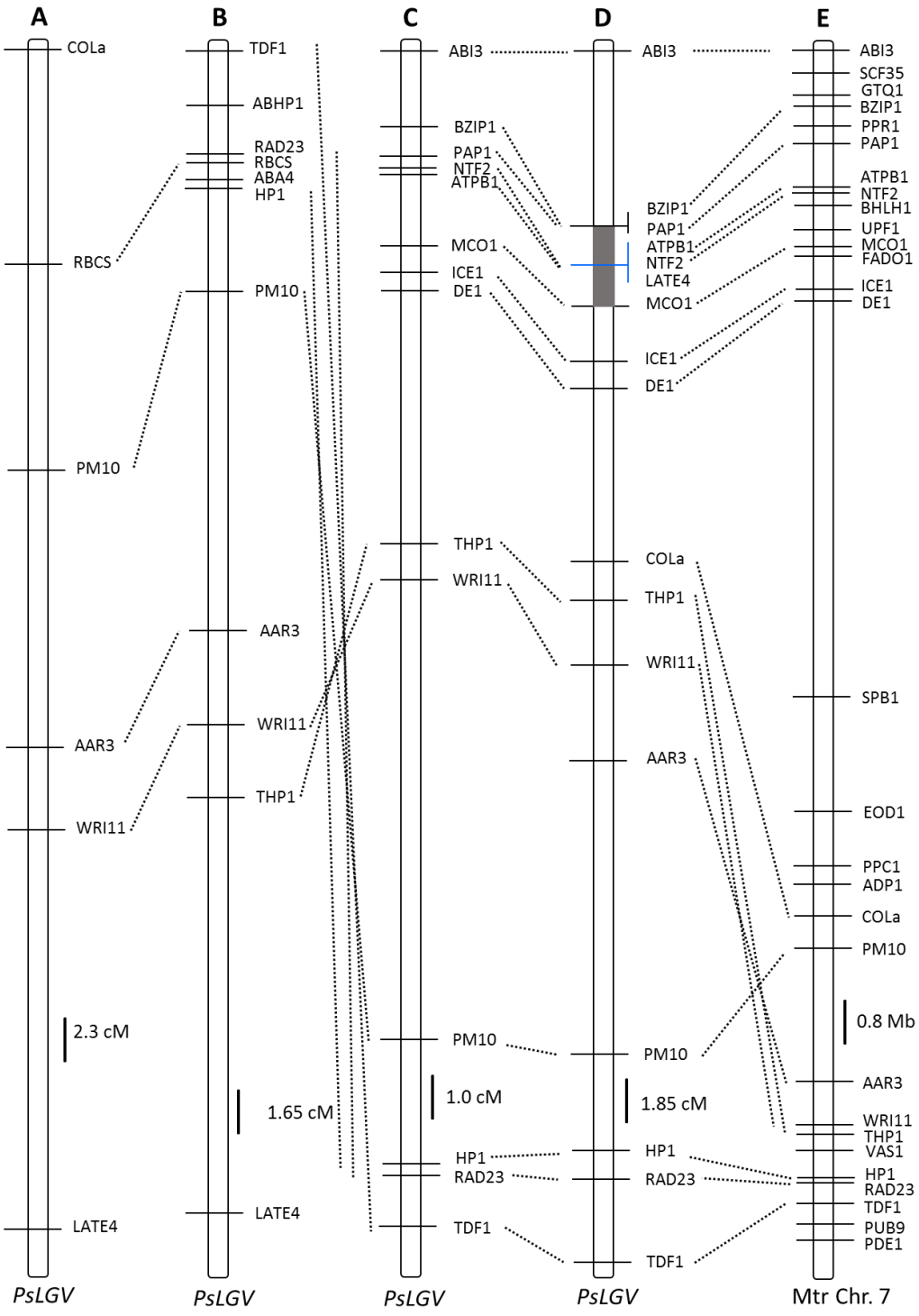


Figure 4.4. Various steps in the development of genetic map for *LATE4* locus. **(A)** Initial map developed by Hecht et al. (unpublished) on *PsLGV* using 89 F_2 individuals of *late4-2* x cv. Terese mapping population. **(B)** Improvement of the existing map conducted in this study by genotyping existing markers in the *late4-1* x cv. Terese F_2 population having 111 individuals. **(C)** Genetic map and order of markers in *PsLGV* developed by Tayeh et al., 2015. **(D)** Fine mapping of *LATE4* locus using a larger *late4-1* X cv. Terese F_2 mapping population having 189 individuals. (**E)** Comparative physical map of *M. truncatula* in chromosome 7 showing position of some of the genes targeted for marker development (see Table 4.2 for all genes). Shaded region represent defined interval showing location of *LATE4* locus between flanking markers.

The newly developed markers (Figure 4.4 B) did not narrow down the map interval significantly even though one marker called *THP1* merely improved the map position which was still located 15.7 cM away from *LATE4* locus. Since most of these new markers were mapped further top of *LATE4* locus in *PsLGV* (Figure 4.4 B, Table 4.2), therefore it was speculated at that stage that *LATE4* locus might be located in the other way, i.e., towards upward direction from *THP1* instead of originally assumed direction towards the bottom of *PsLGV*.

At that time point, Tayeh et al. (2015) published their work on the development of a high resolution, consensus genetic map using freshly innovated Genopea 13.2K SNP array where they mapped 12,802 transcript derived SNP markers through genotyping in 12 pea recombinant inbred lines. This work also delivered a dense syntenic network of pea linkage groups with already sequenced legumes such as *M. truncatula* and it was taken as a basis for further marker development (Figure 4.4 C). Besides, an additional F_2 seed was still available from the *late4-1* x cv. Terese and a supplemental population of 78 individuals was grown to increase the overall population size. Similar to *late3-1*, early and late flowering segregants could be clearly distinguished in this new population based on the scoring of node of flowering initiation (Figure 4.5). In an attempt to remove any potential discrepancy in the order of markers, 56 out of 189 of the F_2 plants that showed WT phenotypes and also had enough

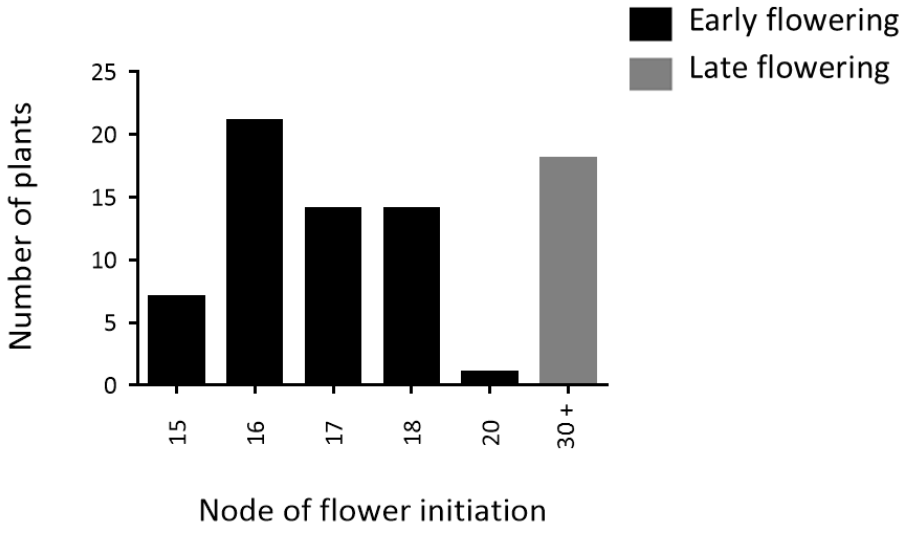
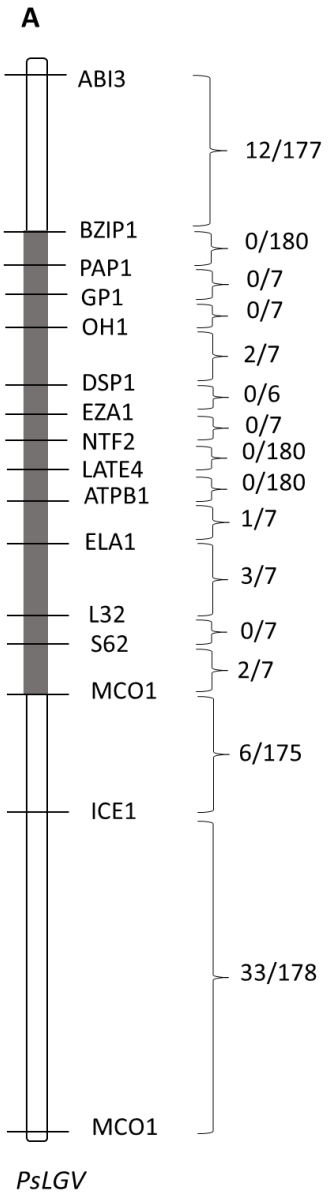


Figure 4.5. Segregation of node of flower initiation among the F_2 plants from *late4-1* X cv. Terese grown under LD condition during the period (2016) of this study.

seeds were grown in the F_3 generation in order to determine whether the genotype of classical marker (*LATE4* locus) for those individuals is homozygous or heterozygous. Using information obtained from these three different approaches, i.e., mapping existing markers in the *late4-1* population, developing new markers based on Tayeh et al. 2015 marker order and F_3 sowing, the position of the *LATE4* locus was remarkably improved which placed it further up of *WRI11* marker in *PsLGV* within genetic distance of 3.2 cM between flanking markers *PAP1* (co-segregated with *BZIP1*) and *MCO1* (Figure 4.4 D, Table 4.2, A1.2). This region corresponded to 2.51 Mb in the physical map of *M. truncatula* chromosome 7 (Figure 4.4 E, Table 4.2).

Further fine mapping was done after the analysis of recombinants as mentioned earlier for *LATE3* in section 4.4.1 which led to the identification of eight important recombinants (Figure 4.6 A-B). Six of these recombinants showed segregation between *LATE4* and *MCO1* while the other two exhibited segregation between *PAP1* and *LATE4* (Figure 4.6 B). Thus, presence of eight recombinants within the interval of 3.2 cM suggested that more markers would help refine the map position of *LATE4*. These additional markers were genotyped mostly on the eight important recombinants apart from the marker *NTF2* and *ATPB1* (Figure 4.6 A-B). This new method of genotyping led to identification of three recombinants, i.e., recombinant 3, 5 and 6 from below which showed segregation between a closer marker namely *L32* and classical marker *LATE4* (Figure 4.6 B, Table 4.2). Another marker namely *ELA1* narrowed down the map position of *LATE4* even further as one of the recombinants, i.e., recombinant 1 showed segregation between these two loci (Figure 4.6 B, Table 4.2). On the other hand, additional recombinant (s) showing segregation between *LATE4* and new markers from the top could not be identified through this process. Nevertheless, the marker *OH1* which showed co-segregation with *PAP1* narrowed down the map position by inference from *M. truncatula* physical position (Figure 4.6 B, Table 4.2). Four markers namely *DSP1*, *EZA1*, *NTF2* and *ATPB1* were found to be co-segregating with *LATE4* locus. The physical distance for the genomic location between *OH1* and *ELA1* in *M. truncatula* chromosome 7 was 0.48 Mb (Table 4.2). The next chapter will thoroughly discuss about the approach of candidate gene selection and analysis for this location.



B

	Recombinant 1	Recombinant 2	Recombinant 3	Recombinant 4	Recombinant 5	Recombinant 6	Recombinant 7	Recombinant 8
ABI3	B	B	A	H	A	H	A	B
BZIP1	B	B	A	H	A	H	A	-
PAP1	B	B	A	H	A	H	A	B
GP1	B	B	A	H	A	H	A	B
OH1	B	B	A	H	A	H	A	B
DSP1	B	B	-	B	A	H	H	B
EZA1	B	B	A	B	A	H	H	B
NTF2	B	B	A	B	A	H	H	B
LATE4	C	B	A	B	A	H	H	B
ATPB1	B	B	A	B	A	H	H	B
ELA1	H	B	A	B	A	H	H	B
L32	H	B	H	B	H	B	H	B
S62	H	B	H	B	H	B	H	B
MCO1	H	H	H	B	H	B	H	H
ICE1	H	H	H	B	H	B	H	H
THP1	A	H	H	B	H	B	H	H

Figure 4.6 Analysis of recombinants that showed segregation between neighbouring markers using genotype data from F_2 mapping population of *late4-1* x cv. Terese. **(A)** Number of recombinants out of total genotyped individuals between neighbouring markers, figure not in scale. **(B)** Genotype of important recombinants used to determine potential location of *LATE4* locus, A = Homozygous *late4-1,* B = Homozygous cv. Terese, H = Heterozygous, C = Homozygous cv. Terese / Heterozygous. Marker names colored with red were genotyped only in the important recombinants.

Table 4.2. Genes attempted for developing markers to map *LATE4* locus using *late4-1* x cv. Terese F2 population in this study. *M. truncatula* Mt4.0v1 database and Pea RNA seq gene atlas database were used for collecting accession numbers. Information on primers and PCR carried out to obtain polymorphism between parental genotypes are given in table A1.2 whereas primers for marker are mentioned in table A1.3.

Locus name	*P. sativum* gene accession number	*M. truncatula* gene accession number	Physical position on *M. truncatula* chromosome 5 (Mb)	Marker type	Marker reference
ABI3	PsCam027818	Medtr7g059330	21.48	CAPS	Aubert et al., 2006
SCF35	PsCam045631	Medtr7g058520	21.09	No marker	-
GTQ1	PsCam048130	Medtr7g057530	20.72	No marker	-
BZIP1	PsCam040141	Medtr7g057160	20.55	HRM	This study
DEAD1	PsCam045248	Medtr7g057250	20.51	No marker	-
PPR1	PsCam000159	Medtr7g056613	20.22	No marker	-
PAP1	PsCam034372	Medtr7g056400	19.89	HRM	This study
OH1	PsCam050951	Medtr7g056233	19.68	HRM	This study
GPI1	PsCam049521	Medtr7g056117	19.53	Size	This study
DSP1	PsCam002046	Medtr7g055923	19.3	Size	This study
ELA1	PsCam010827	Medtr7g055793	19.2	HRM	This study
ATPB1	PsCam033465	Medtr7g055720	19.11	HRM	This study
EZA1	PsCam029108	Medtr7g055660	19.09	HRM	This study
NTF2	PsCam030850	Medtr7g055610	19.04	HRM	This study
BHLH1	PsCam042714	Medtr7g053410	18.8	No marker	-
EAC1	PsCam023290	Medtr7g053290	18.75	No marker	-
L32	PsCam002234	Medtr7g053180	18.7	HRM	This study
S62	PsCam011292	Medtr7g053160	18.7	dCAPS	This study
UPF1	PsCam038395	Medtr7g052250	18.36	No marker	-
MCO1	PsCam050297	Medtr7g051440	18.04	HRM	This study
FADO1	PsCam012946	Medtr7g051240	17.94	No marker	-
ICE1	PsCam045614	Medtr7g451440	17.3	HRM	This study
DE1	PsCam050323	Medtr7g451010	17.11	HRM	This study
SPB1	PsCam012994	Medtr7g028740	9.87	No marker	-
EOD1	PsCam007049	Medtr7g023770	7.77	No marker	-
PPC1	PsCam010674	Medtr7g021530	6.83	No marker	-
ADP1	PsCam035619	Medtr7g020860	6.5	No marker	-
COLa	PsCam010875	Medtr7g018170	5.83	dCAPS	Hecht et al., unpublished
PM10	PsCam001681	Medtr7g016630	5.26	Size	Hecht et al., unpublished
AAR3	PsCam050275	Medtr7g010920	2.8	CAPS	Hecht et al., unpublished
WRI11	PsCam038324	Medtr7g009410	2.05	HRM	Hecht et al., unpublished
MAB1	PsCam014151	Medtr7g009250	1.97	No marker	-

Table 4.2. (continued)

Locus name	*P. sativum* gene accession number	*M. truncatula* gene accession number	Physical position on *M. truncatula* chromosome 5 (Mb)	Marker type	Marker reference
THP1	PsCam010122	Medtr7g009070	1.93	HRM	This study
CAT9	PsCam026869	Medtr7g007850	1.7	No marker	-
VAS1	PsCam040691	Medtr7g007770	1.65	No marker	-
SWEET9	PsCam039934	Medtr7g007490	1.51	No marker	-
ABHP1	PsCam038878	Medtr7g007310	1.43	HRM	This study
ABA4	PsCam001816	Medtr7g007280	1.41	CAPS	This study
MGL1	PsCam009698	Medtr7g007260	1.4	No marker	-
RBCS	PsCam011201	Medtr7g007230	1.39	ASP	Aubert et al., 2006
RBC1	PsCam011201	Medtr7g007120	1.34	No marker	-
SSP1	PsCam057865	Medtr7g007010	1.3	No marker	-
HP1	PsCam059449	Medtr7g407140	1.11	dCAPS	This study
RAD23	PsCam034125	Medtr7g407040	1.06	dCAPS	This study
ATR3	PsCam057504	Medtr7g406970	1.02	No marker	-
NADP1	PsCam033498	Medtr7g406860	0.94	No marker	-
NEF1	PsCam042442	Medtr7g406760	0.86	No marker	-
TDF1	PsCam044291	Medtr7g006560	0.74	HRM	This study
PT1	PsCam029064	Medtr7g006540	0.73	No marker	-
GOT1	PsCam043494	Medtr7g006340	0.65	No marker	-
ENTH1	PsCam048827	Medtr7g006280	0.63	No marker	-
SAMT1	PsCam023112	Medtr7g006060	0.5	No marker	-
HMB1	PsCam057867	Medtr7g005970	0.43	No marker	-
GS1	PsCam035287	Medtr7g005950	0.42	No marker	-
PUB9	PsCam011495	Medtr7g005940	0.4	No marker	-
SPP1	PsCam023099	Medtr7g405780	0.26	No marker	-
LTP1	PsCam039604	Medtr7g405770	0.24	No marker	-
SAG29	PsCam011330	Medtr7g405730	0.23	No marker	-
HPG1	PsCam045353	Medtr7g005390	0.09	No marker	-
PDE1	PsCam014151	Medtr7g005380	0.08	No marker	-

4.5 Discussion

4.5.1 Translational genomics is an effective tool in fine mapping pea mutant loci such as late3 and late4

In line with the strategy mentioned in section 4.1.2, Hecht et al (unpublished) started the process of genetic map development for *LATE3* locus by using the phenotypic data generated for 192 F_2 individuals through a cross between cv. Terese and *late3-1* genotypes (Weller et al unpublished). This initial map which placed the *LATE3* locus at the middle of *PsLGIII* has been improved significantly during the current study by reanalysing old markers in the existing and a new, larger F_2 mapping population as well as by genotyping new markers (Figure 4.1 A-D). Analysis of the number of recombinants between neighbouring markers identified eight important recombinants between the flanking markers *BTB1* and *SPS1* along with *LATE3* locus co-segregating the marker *APP2* (Figure 4.3 A-B). Order of all the markers in *PsLGIII* were highly conserved (co-linear) to *M. truncatula* chromosome 3 (Figure 4.1 C-D) which was also observed during similar studies conducted previously for other pea late flowering mutant loci (Hecht et al. 2007; Sussmilch 2014; Ridge et al. 2016). This interval in chromosome 3 of *M. truncatula* has 62 genes and identification of potential candidates from these genes is discussed in chapter 5.

On the other hand, the genetic map for the *LATE4* locus was at more preliminary stage than the *LATE3* locus at the commencement of this project. The initial map generated by Hecht et al. (unpublished) which positioned the *LATE4* locus at the base of the *PsLGV* did not have markers on both side, rather only on the top side of the map with a distance of >20 cM. The initial perception of potential locality of *LATE4* locus downstream of the nearest marker *WRI11* rather than upstream was proved wrong during the period of this study by developing a fine map using the marker order of the consensus map published by Tayeh et al., 2015 as a reference (Figure 4.4 A-E). Generation of F_3 plants from a large number of F_2 plants also contributed largely towards the improvement of the map position of *LATE4* locus. Analysis of the number of recombinants between neighbouring markers identified eight important recombinants between a larger region bordered by markers *S62* and *OH1* with four markers co-segregating with the *LATE4* locus.

Comparison of the genetic map of *PsLGV* developed in the present study with corresponding physical map of *M. truncatula* chromosome 7 (Figure 4.4 D-E) revealed that order of all the molecular markers apart from *PM10*, *AAR3*, *WRI11* and *THP1* were syntenic to orthologous *M. truncatula* genes which suggest potential chromosomal reversion for the region between *PM10* to *THP1* in pea (3.33 Mb in *M. truncatula* chromosome 7, Table 4.2). This chromosomal rearrangement within the aforementioned region is the likely reason because of which it was initially assumed that *LATE4* locus is located further down than *WRI11* marker in *PsLGV*. However, comparative genetic map of *PsLGV* from the current study and that of Tayeh et al., 2015 showed that order of all the markers from both the maps are consistent. Therefore, this high resolution consensus map from Tayeh et al., 2015 could also be used as a reference in future studies as chromosomal rearrangement between pea and *M. truncatula* for different linkage groups such as *PsLGIII* (syntenic to *M. truncatula* chromosome 2 and 3), *PsLGVI* (syntenic to *M. truncatula* chromosome 2 and 6) is known (Kaló et al. 2004). There are 54 genes located in the syntenic region of chromosome 7 of *M. truncatula* between markers *S62* and *OH1* in *PsLGV* and analysis of likely candidate genes is carried out in next chapter.

Availability of co-segregating markers with *LATE3* and *LATE4* loci *PsLGIII* and *PsLGV* respectively provide strong indication that both the candidate genes are likely to be located within the defined interval. Besides, both the maps generated in the present study further strengthened current understanding about the applicability of *M. truncatula* as a tool of translational genomics for mapping developmental genes in pea. In other words, high degree of evolutionary conservation of the order of orthologous genes (co-linearity) within macro and microsyntenic regions (apart from *PM10-THP1* in *PsLGV*) as revealed in the current study between pea genetic map and corresponding *M. truncatula* physical map (Figure 4.1 C-D, 4.4 D-E) is in line with already available information (Sanders et al. 2011). Thus, the present study established the reliability of using *M. truncatula* for future purposes even when the pea genome sequence would be publicly accessible.

4.5.2 Concluding remark

The genetic maps generated in this chapter laid down a crucial foundation for candidate gene analysis and molecular characterization of the *LATE3* and *LATE4* genes. For instance, co-segregation between the mutant loci and other genetic markers provides strong indication that the underlying *LATE3* and *LATE4* genes are likely to be located within the defined genetic

intervals in *PsLGIII* and *PsLGV* respectively. Besides, higher degree of synteny within the region of interest between pea and *M. truncatula* that has been unveiled in this chapter led to further speculation that functional annotation and physical position of genes in corresponding *M. truncatula* chromosomes is likely to be equally important for ultimate determination of *LATE3* and *LATE4* gene identity as it was for developing the genetic maps in pea. Moreover, marker information established for *PsLGIII* and *PsLGV* would serve as a useful resource in future for pea as well as legume community involving molecular genetics research within these chromosomal locations.

Chapter 5: Candidate gene analysis for *LATE3* and *LATE4* loci

5.1 Introduction

5.1.1 Candidate based gene identification

A major goal in the field of plant genetics involving molecular characterization of mutants is to unveil the identity of the gene function of which has been impaired and resulted in aberrant phenotype. Once a mutant locus has been mapped with a narrow genetic interval, then positional candidate based gene identification is a suitable follow-up strategy. Three subsequent steps are involved in the analysis of candidate genes (Pflieger et al. 2001). These are as follows: (i) candidate genes could be chosen based on the molecular and physiological studies which are termed as functional candidates. On the other hand, all the neighbouring genes resembling close linkage with the mutant locus could potentially be considered as positional candidate genes. (ii) molecular polymorphism needs to be detected between the parents which can then be exploited to genotype the entire F_2 population or only important recombinants, thereby establish co-segregation with mutant loci (phenotypic marker) and possibly with nearest genetic markers as well. (iii) once the previous two criteria are fulfilled, then relevant transgenic complementation studies should be carried out (if possible) in order to validate involvement of the candidate gene with phenotypic variation. For instance, full length cDNA of the WT could be inserted into loss-of-function mutants and thereby try to establish function of WT protein and recover the WT phenotype. However, in case of gain-of-function mutants, transgenic phenocopy experimentation is rather suitable where the full length WT cDNA is overexpressed in the WT background and resultant phenotype is compared with transgenic lines generated in the WT background by inserting the mutated version of the cDNA (Exner et al. 2010). Such experimentation could assist in determining whether the gain-of-function mutant derived due to elevated level of expression of the protein of interest. Nevertheless, if multiple independent mutant alleles are available for the locus under investigation, then identification of functional polymorphism in the putative candidate gene for each of the mutant alleles could mean that there is causal relationship between the observed mutant phenotype and the candidate gene (Weller et al. 2013). In such case, transgenic complementation/phenocopy studies would not be necessary.

The model species, *A. thaliana* serves as an excellent tool in the overall process of molecular characterization of mutants. It has already been shown that a number of flowering related genes from *A. thaliana* are conserved in pea and *M. truncatula* (Hecht et al. 2005). Comparative analysis of genetic and physical maps between closely related species could also assist in choosing candidate genes which is not needed when high quality genome sequences of the species under investigation is available. Due to the availability of genomics sequences of *M. truncatula* (Tang et al. 2014), transcriptome database of pea (Alves-Carvalho et al. 2015) as well as genetic maps (Kaló et al. 2004; Bordat et al. 2011; Sindhu et al. 2014; Tayeh et al. 2015a) with information on comparative positions of functional genes between these two closely related species boosted identification of flowering time genes in pea. Phylogenetic analysis of the identified gene could provide direction about the potential evolutional history of the gene under investigation whereas gene expression data from geneatlas assist to generate understanding about the potential regulatory role of the gene.

A limiting factor for comparative genomics research in pea is that the *M. truncatula* genome sequence is still not entirely complete due to error in the orientation of BAC contigs causing presence of gaps in the assembly (Tang et al. 2014). Often it might be required to increase map resolution for the defined location in the case where obvious candidate genes are not found or potential candidate genes did not have causal mutations. For this, development of high resolution map by generating larger population or adding more markers is the relevant step to pinpoint upon an even narrower genetic interval in the quest for candidate gene identification.

In the previous chapter, high resolution genetic map for both *LATE3* and *LATE4* loci in pea were developed by using information of physical map of *M. truncatula*. The defined interval for *LATE3* between markers *BTB1* and *SPS1* in *PsLGIII* and *LATE4* between markers *S62* and *OH1* in *PsLGV* formed the basis of candidate gene analysis which have been carried out in the present chapter.

5.1.2 Exploitation of RNA sequencing in candidate gene selection

Selection of candidate genes within the defined interval through comparative genomics could involve direct sequencing of these genes if orthologues of such genes are present in the organism of interest. This procedure is potentially time consuming. The advent of next

generation sequencing (NGS) technologies have revolutionized molecular genetics based research which can be used to generate millions of reads in a cost-effective, high-throughput manner for addressing various research questions. A number of whole genome sequencing (WGS) have been exploited in this regard both in animals, e.g. zebra fish and plants, e.g. *A. thaliana* for which entire genome sequence is available (Bowen et al. 2012; Leshchiner et al. 2012; Obholzer et al. 2012; Thole and Strader 2015).

However, WGS is an expensive process and not all species have their entire genome sequenced. RNA sequencing has several benefits for NGS mapping in comparison to WGS with a reduced cost (Miller et al. 2013; Garg and Jain 2013; Schneeberger 2014). These are as follows: i. the amount of sequencing required for high coverage and high quality data is reduced, ii. gene based markers could be developed, iii. effect of the mutation on global gene expression could be potentially determined, iv., impact of the candidate mutant on RNA splicing can be studied, v. useful for organisms with no available genome sequence database.

Various important aspects that needs to be taken into consideration for performing RNA sequencing in plants have been described by Garg and Jain, 2013. Two main factors are choice of appropriate tissue and sequencing depth. Wide range of tissue samples represented by different stages of the lifecycle of the plant could be pooled for transcript construction for most of the transcripts expressed in that species. In contrast, sequencing individual samples from specific tissues or growth stages could potentially be used for both transcript construction as well as gene expression analysis. Similarly, cost-effective, low sequencing depth is considered sufficient for getting sequence of most of the transcriptomes and could be used for assembling against an available reference genome. However, some crucial genes might have low expression or the reference sequence might not be available for some species. In such case, higher sequencing depth might be needed for getting enough reads for certain transcripts or to carry out de novo assembly of the transcriptome. Apart from these, RNA sequencing is limited by the fact that mutations occurring in introns and regulatory sequences cannot be identified through this process.

Previous functional characterization of flowering mutants in pea by Weller et al. involved direct sequencing of some obvious candidate genes to find the causal mutations (Hecht et al. 2007; Liew et al. 2009; Hecht et al. 2011; Liew et al. 2014; Sussmilch et al. 2015; Ridge et al.

2016). RNA sequencing platform was recently established for expression studies in pea (Ridge et al. 2016) and during the course of this project, it became also possible to explore the use of this technology for identification of causal mutations. Usage of such approach was considered to be useful specifically for cases where clear functional/positional candidate genes similar to previous studies may not be available. Therefore, RNA sequencing has been exploited in the present study involving *late3* and *late4* mutants for the detection of functional polymorphisms in the transcripts of candidate genes located within the defined interval of *PsLGIII* and *PsLGV* respectively.

5.2 Chapter aim

The general aim of this chapter was to evaluate candidate genes for *LATE3* and *LATE4* loci. To this end, analysis of few functional and positional candidate genes have been carried out through direct sequencing and relevant phylogenetic analysis. This was followed by analysis of RNA sequencing data of all the pea transcripts identified through comparative mapping between pea and *M. truncatula* within the defined intervals for both the mutant loci. Once the *LATE3* and *LATE4* genes were identified, it was essential to carry out further studies to validate the nature of mutations, hence experiments regarding alternative splicing were conducted. Finally, in order to strengthen understanding about the role of the *LATE4* gene and the degree to which its function might be conserved in *A. thaliana,* phenotypic comparison of *late4* mutants with the respective T-DNA insertion mutant of orthologous *A. thaliana* gene was performed.

5.3 Materials and methods

Various materials and methods used specifically for this chapter are mentioned here. The general materials and methods exploited in this chapter are given in chapter 2. Details of sequences and alignments are provided in appendix 2.

5.3.1 Annotated function of orthologous genes through comparative genomics

Accession numbers for genes located within the genetic map position bordered by *BTB1* and *SPS1* in *PsLGIII* and between *OH1* and *S62* in *PsLGV* were collected from comparative physical map position in *M. truncatula* genome database (*M. truncatula* Mt4.0v1, https://phytozome.jgi.doe.gov/pz/portal.html#). Physical position and annotation of the genes within the aforementioned locations were compiled in order to determine the presence of any functional candidate genes. Respective accession numbers for pea transcripts were then collected by BLASTn of the *M. truncatula* transcript against the pea RNA sequence gene atlas database (http://bios.dijon.inra.fr/FATAL/cgi/pscam.cgi). Only one gene accession number was collected for *M. truncatula* genes where multiple transcript variant information is available in the database. Refer to table 5.1 and 5.2 for these data. All these information assisted in the determination and selection of few functional and positional candidates that were carried out in this chapter.

5.3.2 RNA sequencing

RNA preparation and RNA sequencing were performed by Dr. Valerie Hecht and reads were provided for analysis. These reads were derived from two different RNA sequencing runs, with an initial low coverage run supplemented by higher coverage. In the first run, NGB5839 (WT), *late3-1* and *late4-1* genotypes were used. For the second run, *late3-1* and isogenic *late5* mutant as effective WT was used for comparison. Both the runs were carried out using following tissue samples: embryo from two days imbibed seeds and newly expanded leaves and apex from four weeks old plants.

Quality of the reads generated from Illumina RNA sequencing machine were assessed in FASTQC which is a web based platform in galaxy (Giardine et al. 2005). Quality scores of >28 across all bases of the reads as determined by FASTQC were considered of good quality. Reads were then imported in genious 8.0.4 software (Kearse et al. 2012) and paired end from the two relevant files for each genotypes were created with a distance of 100 bp between them.

5.3.3 Phylogenetic analysis

Amino acid sequences collected for a particular gene from *A. thaliana* were used for blastP search of orthologous proteins in pea and *M. truncatula* and these were aligned in geneious 8.0.4 software (Kearse et al. 2012) using ClustalW alignment (Higgins et al. 1996; Thompson et al. 1997; Larkin et al. 2007) upon selection of BLOSUM matrix (Henikoff and Henikoff 1992; Eddy 2004). After that, the aligned protein sequences were used for creating phylogenetic tree using PAUP* 4.0b10 by choosing maximum parsimony method and 1000 replications of bootstrapping (Swofford 2001).

5.3.4 Open source gene expression profile

Gene expression profile for the candidate genes *Cyclin dependent kinase 8* (*CDK8*) and *Cyclin C* (*CYCC1*) were collected for different developmental organs for *A. thaliana* from AtGenExpress data (http://jsp.weigelworld.org/expviz/expviz.jsp), *O. sativa* from Rice Expression Database - IC4R (http://expression.ic4r.org/), *Lotus* from *Lotus japonicus* gene expression atlas (https://ljgea.noble.org/v2), *Medicago* from *M. truncatula* gene expression atlas (https://mtgea.noble.org/v2) and pea from RNA-seq gene atlas (http://bios.dijon.inra.fr/FATAL/cgi/pscam.cgi).

5.3.5 Direct sequencing of candidate genes

Putative SNPs detected in any positional candidate genes were further validated by direct sequencing of cDNA and/or gDNA of the relevant gene. Information on the primer sequences are provided in table A1.3.

5.3.6 Experiment involving T-DNA insertion lines for A. thaliana CYCC1 gene

5.3.6.1 Selection of lines

T-DNA insertion mutants for *A. thaliana CYCC1.1* gene (AT5G48640) *CYCC1.2* (AT5G48630) were chosen using the T-DNA express - *A. thaliana* gene mapping tool (http://signal.salk.edu/cgi-bin/tdnaexpress) from the SALK institute genomic analysis laboratory (O'Malley et al. 2015). Among the available lines, two lines SALK_053291C and SALK_039400C named in this study as *cycc1-1* and *cycc1-2* respectively were chosen for the current study as these have been mentioned as confirmed homozygous knockout mutants in the background of Columbia-0. Seeds from these two lines were ordered. WT, Columbia-0

seeds were provided by Dr. Lu Wang from the School of Natural Sciences, University of Tasmania (personal communication).

5.3.6.2 Sowing and growth conditions

Before sowing, seeds were surface sterilized with 30% bleach + 0.1% Triton X-100 for 5-10 minutes. Bleach was then removed by rinsing the seeds 5-7 times with sterile distilled water. Approximately 20 seeds from each genotype were then placed on a filter paper soaked with sterile distilled water in a petridish and incubated at 4° C for three days in order to increase the rate of germination. After that, one seed was placed on top of the soil of a small pot having dimension of approximately 6 cm X 4 cm. These small pots were placed in a large tray and shifted to a growth chamber cabinet with 16 hours of photoperiod, 20° C temperature, with average light intensity of 120-150 µmol. Initially, pots were covered with a lid to keep humidity which was removed after 25 days of growth. Rosette leaves from each genotypes were harvested in two replications after 40 days of growth for DNA extraction and genotyping of the plants.

5.3.6.3 Genotyping

DNA was extracted from WT and putative mutants and concentrations were measured using the same protocol mentioned in chapter 2.

For each of the above mentioned SALK lines, gene specific LP and RP primers were designed using the default settings of the T-DNA primer design tool (http://signal.salk.edu/tdnaprimers.2.html) from SALK institute genomic analysis laboratory website. The T-DNA insertion specific LB primer was chosen from the website as recommended by them. Primer sequences for these primers are mentioned in table A1.4.

5.3.6.4 Phenotypic data collection

Flowering initiation in the insertion mutants and WT genotype were recorded by counting the number of rosette leaves at the onset of flower.

Length of 1st and 10th silique as well as the length of pedicle of these two siliques were measured as phenotypic variation for these traits between the WT and insertion mutants were observed in this study.

Plant height and total branching were measured at the time of harvest 90 days after sowing. Total branching was calculated as a ratio of total lateral branch length and total stem length.

Total number of siliques formed along the main stem were also counted at the time of harvest 90 days after sowing.

5.4 Results

5.4.1 Evaluation of SDG10 and SDG15 as candidates for LATE3 and LATE4 respectively

Prior to the start of candidate gene analysis for *LATE3* and *LATE4* loci within the defined interval at *PsLGIII* and *PsLGV* respectively, information on the annotated functions of the *M. truncatula* genes located at syntenic physical map position was collected and their corresponding pea orthologues were identified via blastP search against the pea transcriptome database. There were 62 genes within *BTB1* and *AAP2* markers in chromosome 3 of *M. truncatula* (Table 5.1) whereas this number was 54 for the region between *S62* and *OH1* markers on chromosome 7 of *M. truncatula* (Table 5.2).

The next phase of candidate gene analysis relied on screening for functional/positional candidates within the syntenic region of *M. truncatula* chromosome 3 and 7 respectively that are known to be involved or have some possible link to flowering time mechanism. Such screening led to notifying the presence of *MtEZA1* (Medtr1g086980) at 19.09 Mb of *M. truncatula* chromosome 7 which is an orthologue of *A. thaliana SET domain gene 10* (*AtSDG10*) or *SWINGER* (*AtSWN1*) or *ENHANCER OF ZESTE* (*AtEZA1*). The pea orthologue of this gene, *PsEZA1* (PsCam029108) showed co-segregation with *LATE4* locus when genotyped in important recombinants of *late4-1* x cv. Terese F_2 population (see Figure 4.6 B). In *A. thaliana*, EZA1 acts as H3K27 methytransferase which is a component of the polycomb repressive complex 2 (PRC2) and it works in concert with some other proteins in order to promote accumulation of H3K27me3 repressive epigenetic marks at the *FT* gene locus (Jiang et al. 2008; Farrona et al. 2011). Besides, it interacts with other polycomb group proteins in order to repress FLC (Wood et al. 2006).

Based on the similarities in the phenotype of *late3* and *late4* mutants, it was hypothesized in chapter 3 that *LATE3* and *LATE4* might have similar functions and they may act in the same regulatory pathway. The *MtSDG1//MtATXR5* (Medtr3g095840) gene located at 43.79 Mb of *M. truncatula* chromosome 3 which is orthologous to *A. thaliana SET domain gene 15* (*AtSDG15*)/*Trithorax related gene ATXR5* was identified simultaneously which has similarities in function to that of aforementioned *AtSDG10* gene. Trithorax related proteins are H3K4 methyltransferases and are known to act in a protein complex to enhance deposition of gene

Table 5.1. Orthologous genes located in syntenic *M. truncatula* chromosome 3 for candidate region determined through flanking markers *BTB1* and *SPS1* in *P. sativum LGIII*. *M. truncatula* gene ID have been taken from *Medicago truncatula* Mt4.0v1 database (https://phytozome.jgi.doe.gov/pz/portal.html#) and *P. sativum* accession numbers have been collected through blastP search using *M. truncatula* protein sequences against Pea RNA seq gene atlas database (http://bios.dijon.inra.fr/FATAL/cgi/pscam.cgi)

M. truncatula gene ID	Annotation in *M. truncatula* genome database	*M. truncatula* physical position	*P. sativum* transcript ID
Medtr3g096160	BTB/POZ domain plant protein	Chr3:43945334-43942445	PsCam000088
Medtr3g096190	Nucleolar essential protein	Chr3:43952355-43948147	PsCam034774
Medtr3g096200	Glucan synthase-like protein	Chr3:43969447-43952945	PsCam035309
Medtr3g096230	MAP kinase kinase kinase	Chr3:43980361-43985582	PsCam048280
Medtr3g096240	DNA repair metallo-beta-lactamase	Chr3:43986916-43990495	PsCam000112
Medtr3g096290	Cyanobacterial and plant NDH-1 subunit O	Chr3:44005926-44004783	PsCam001940
Medtr3g096300	ATP-binding ABC transporter	Chr3:44008517-44012720	PsCam034645
Medtr3g096310	GDP-mannose transporter GONST1	Chr3:44013420-44018709	PsCam030881
Medtr3g096320	GDP-mannose transporter, putative	Chr3:44019323-44021064	PsCam052035
Medtr3g096330	Glycoside hydrolase family 28 protein	Chr3:44026955-44022720	PsCam037303
Medtr3g096350	Small GTPase family RAB protein	Chr3:44038043-44043862	PsCam029411
Medtr3g096370	E3 ubiquitin ligase PUB14	Chr3:44050741-44048824	PsCam050053
Medtr3g096380	Cation/H+ exchanger 3	Chr3:44056812-44054012	PsCam059372
Medtr3g096390	2-oxoisovalerate dehydrogenase E1 component, alpha subunit	Chr3:44061897-44058501	PsCam049851
Medtr3g096400	NAC transcription factor-like protein	Chr3:44063892-44061980	PsCam039877
Medtr3g096480	Methyltransferase PMT26-like protein	Chr3:44098927-44104608	PsCam032493
Medtr3g096500	Gibberellin 20-oxidase	Chr3:44115170-44112343	PsCam046205
Medtr3g096520	E3 ubiquitin-protein ligase RGLG2-like	Chr3:44128294-44132918	PsCam004937
Medtr3g096525	DNA polymerase delta catalytic subunit	Chr3:44150897-44133139	PsCam001809
Medtr3g096555	Adenine nucleotide alpha hydrolase-like domain kinase	Chr3:44157839-44153991	PsCam008813
Medtr3g096562	FAD/FMN-binding oxidoreductase	Chr3:44172767-44169527	PsCam033826
Medtr3g096568	Hypothetical protein	Chr3:44189094-44187154	PsCam001401
Medtr3g096590	Tyrosine kinase family protein	Chr3:44199073-44193476	PsCam054769
Medtr3g096670	F-box plant protein, putative	Chr3:44234632-44238356	PsCam034500
Medtr3g096675	Farnesylcysteine lyase	Chr3:44239339-44242962	PsCam037090
Medtr3g096700	Ribonuclease	Chr3:44253252..44257088	PsCam046150
Medtr3g096710	NAD-dependent protein deacetylase SRT2	Chr3:44265774-44259211	PsCam000535
Medtr3g096750	PsbP domain protein	Chr3:44277737-44279975	PsCam039482
Medtr3g096760	Amino acid permease	Chr3:44281053-44284896	PsCam012934
Medtr3g096780	Chromosome condensation regulator RCC1 repeat protein	Chr3:44289276-44296458	PsCam050308
Medtr3g096790	DnaJ domain protein	Chr3:44297240-44304262	PsCam042463
Medtr3g096830	Transmembrane amino acid transporter	Chr3:44319082-44314161	PsCam049915
Medtr3g096860	DUF3727 family protein	Chr3:44329588-44334628	PsCam021158
Medtr3g096870	Neutral alpha-glucosidase	Chr3:44335718-44341726	PsCam049087
Medtr3g096890	Transmembrane protein, putative	Chr3:44344587-44342621	PsCam039677
Medtr3g096900	Beta-galactosidase	Chr3:44353075-44345837	PsCam042665
Medtr3g096920	NAC-like transcription factor	Chr3:44370629-44372461	PsCam048115
Medtr3g096930	Eukaryotic aspartyl protease family	Chr3:44376543-44372751	PsCam048369
Medtr3g096940	Hypothetical protein	Chr3:44380595-44383167	PsCam037556
Medtr3g096960	Cyclin-dependent kinase	Chr3:44404917-44393350	PsCam048317
Medtr3g096980	Fatty acid amide hydrolase-like protein	Chr3:44418973-44424310	PsCam035372

(continued next page)

Table 5.1. (continued)

M. truncatula gene ID	Annotation in *M. truncatula* genome database	*M. truncatula* physical position	*P. sativum* transcript ID
Medtr3g096990	Filament-like plant protein	Chr3:44428717-44424567	PsCam035385
Medtr3g097010	CBS/octicosapeptideBemp1 (PB1) domain	Chr3:44432172-44439603	PsCam042743
Medtr3g097015	Trafficking protein particle complex	Chr3:44443782-44442988	PsCam011333
Medtr3g097060	Chromosome condensation regulator RCC1 repeat protein	Chr3:44474667-44464496	PsCam023312
Medtr3g097090	Nucleotidyltransferase family protein	Chr3:44490299-44481962	PsCam034443
Medtr3g097170	Hypothetical protein	Chr3:44539797-44536388	PsCam039622
Medtr3g097220	Magnesium transporter N	Chr3:44559050..44564779	PsCam010752
Medtr3g097240	DUF668 family protein	Chr3:44568853-44578000	PsCam042642
Medtr3g097260	Phosphatidylinositol-4-phosphate 5-kinase	Chr3:44600133-44594979	PsCam034424
Medtr3g097270	OSBP(oxysterol-binding protein)-related	Chr3:44607533-44603505	PsCam059306
Medtr3g097280	ENTH/VHS/GAT family protein	Chr3:44612030-44618283	PsCam037057
Medtr3g097290	Ferrochelatase	Chr3:44624046-44618822	PsCam034607
Medtr3g097320	Sucrose non-fermenting-like kinase 2	Chr3:44634053-44630416	PsCam050549
Medtr3g097350	Patellin-3 protein	Chr3:44640940-44644067	PsCam036496
Medtr3g097450	Myb transcription factor	Chr3:44673726-44675570	PsCam001579
Medtr3g097560	Cation-transporting ATPase, putative	Chr3:44714904-44728896	PsCam042403
Medtr3g097570	DYAD-like protein	Chr3:44730358-44733848	PsCam049636
Medtr3g098060	Protein prenylyltransferase superfamily	Chr3:44747591-44736028	PsCam035950
Medtr3g098070	CDPK-related kinase	Chr3:44756354-44761409	PsCam033289
Medtr3g098130	Acyl-CoA thioesterase, putative	Chr3:44787465-44785690	PsCam046255
Medtr3g098240	Suppressor OF protein silencing protein	Chr3:44840494-44835486	PsCam045416

Table 5.2. Orthologous genes located in syntenic *M. truncatula* chromosome 7 for candidate region determined through markers *S62* and *OH1* in *P. sativum LGV*. *M. truncatula* gene ID have been taken from *Medicago truncatula* Mt4.0v1 database (https://phytozome.jgi.doe.gov/pz/portal.html#) and *P. sativum* accession numbers have been collected through blastP search using *M. truncatula* protein sequences against Pea RNA seq gene atlas database (http://bios.dijon.inra.fr/FATAL/cgi/pscam.cgi).

M. truncatula gene ID	Annotation in *M. truncatula* genome database	*M. truncatula* physical position	*P. sativum transcript ID*
Medtr7g053160	40S ribosomal protein S6-2	Chr7:18698849-18702010	PsCam011292
Medtr7g053180	60S ribosomal protein L32-1	Chr7:18703682-18705988	PsCam002234
Medtr7g053200	Serine/Threonine-kinase OXI1-like	Chr7:18709700-18707718	PsCam038370
Medtr7g053220	Histidine triad nucleotide-binding	Chr7:18713828-18718572	PsCam021303
Medtr7g053230	Hypothetical protein	Chr7:18719707-18719113	PsCam026961
Medtr7g053240	pentatricopeptide (PPR) repeat protein	Chr7:18723182-18719939	PsCam035812
Medtr7g053260	Plant U-box protein	Chr7:18729058-18725963	PsCam050024
Medtr7g053290	Enoyl-acyl-carrier reductase	Chr7:18751674-18744136	PsCam023290
Medtr7g053310	DUF506 family protein	Chr7:18758577-18757455	PsCam039273
Medtr7g053330	Processing peptidase	Chr7:18766364-18760054	PsCam042823
Medtr7g053360	Pectinesterase/pectinesterase inhibitor	Chr7:18786395-18780789	PsCam037771
Medtr7g053410	BHLH transcription factor	Chr7:18808900-18812617	PsCam042714
Medtr7g053450	40S ribosomal S8-like protein	Chr7:18838026-18834797	PsCam021339
Medtr7g053460	Plant EC metallothionein family 15	Chr7:18843119-18842766	PsCam043919
Medtr7g053470	FKBP-type peptidyl-prolyl cis-trans isomerase	Chr7:18851017-18844062	PsCam050614

(continued next page)

Table 5.2. (continued)

M. truncatula gene ID	Annotation in *M. truncatula* genome database	*M. truncatula* physical position	*P. sativum* transcript ID
Medtr7g053480	Hypothetical protein	Chr7:18852111-18851761	PsCam012869
Medtr7g053500	Glucan endo-1,3-beta-glucosidase-like	Chr7:18856826-18859936	PsCam045678
Medtr7g053520	50S ribosomal protein L7/L12	Chr7:18863632-18861167	PsCam001210
Medtr7g053620	Disease resistance protein (TIR-NBS-LRR class)	Chr7:18923709-18924910	PsCam035395
Medtr7g055590	Transducin family protein/WD-40 repeat	Chr7:19027415-19024828	PsCam000439
Medtr7g055600	Glycosyltransferase Quasimodo protein, putative	Chr7:19029384-19031753	PsCam048130
Medtr7g055610	Nuclear transport factor 2 and RNA recognition motif protein	Chr7:19043619-19039158	PsCam030850
Medtr7g055620	Raffinose synthase or seed inhibition protein Sip1	Chr7:19052899-19057983	PsCam036022
Medtr7g055630	Ankyrin repeat protein	Chr7:19059709-19058453	PsCam029431
Medtr7g055655	Hypothetical protein	Chr7:19080552-19079074	PsCam035904
Medtr7g055650	Amine-terminal domain cyclin	Chr7:19074172-19082353	PsCam050605
Medtr7g055660	Histone-lysine N-methyltransferase	Chr7:19084301-19092927	PsCam029108
Medtr7g055700	GTP cyclohydrolase	Chr7:19097839-19101971	PsCam037210
Medtr7g055710	UDP-glucosyltransferase family protein	Chr7:19104244-19102312	PsCam024790
Medtr7g055720	ATP-binding protein, putative	Chr7:19113432-19108147	PsCam033465
Medtr7g055743	Quinone-oxidoreductase-like protein	Chr7:19132762-19128430	PsCam017510
Medtr7g055747	Transmembrane protein, putative	Chr7:19135802-19134084	PsCam000368
Medtr7g055793	Cytochrome P450 family protein	Chr7:19205156-19208490	PsCam010827
Medtr7g055807	Wound-inducible basic family protein	Chr7:19215197-19212691	PsCam043857
Medtr7g055857	Ethylene response factor	Chr7:19252386-19257484	PsCam029356
Medtr7g055863	Transcriptional activator FHA1	Chr7:19263207-19260879	PsCam007040
Medtr7g055917	Leguminosin group486 secreted peptide	Chr7:19297533-19297994	PsCam052326
Medtr7g055923	Dual specificity phosphatase-like protein	Chr7:19300670-19299724	PsCam002046
Medtr7g055953	F-box protein interaction domain protein	Chr7:19326829-19325757	PsCam045822
Medtr7g056020	F-box/RNI/FBD-like domain protein	Chr7:19371398-19367647	PsCam056096
Medtr7g056023	F-box protein interaction domain protein	Chr7:19373868-19374792	PsCam038127
Medtr7g056030	Nodule Cysteine-Rich (NCR) secreted peptide	Chr7:19379049-19378558	PsCam055630
Medtr7g056073	Basic helix loop helix protein, putative	Chr7:19429810-19428206	PsCam033914
Medtr7g056117	GPI transamidase component PIG-S, related protein	Chr7:19531472-19524387	PsCam049521
Medtr7g056133	B-cell receptor-associated-like protein	Chr7:19556916-19546881	PsCam038311
Medtr7g056147	RPM1 interacting protein 4 transcript	Chr7:19567302-19571423	PsCam036246
Medtr7g056153	Cell division control-like protein	Chr7:19577711-19575486	PsCam045481
Medtr7g056170	F-box and associated interaction domain	Chr7:19600144-19597436	PsCam023210
Medtr7g056173	F-box protein interaction domain protein	Chr7:19605163-19603928	PsCam056707
Medtr7g056183	RING/U-box protein	Chr7:19624899-19620839	PsCam049679
Medtr7g056197	Double Clp-N motif P-loop nucleoside triphosphate hydrolase superfamily protein, putative	Chr7:19636346-19633671	PsCam042433
Medtr7g056207	CXE carboxylesterase	Chr7:19643351-19642281	PsCam014058
Medtr7g056233	Oleoyl-[acyl-carrier] hydrolase (acyl-ACP thioesterase)	Chr7:19680783-19673229	PsCam050951

function promoting epigenetic mark H3K4me3 at the 5'UTR of the *FLC* gene which then represses the function of key flowering gene *FT* in *A. thaliana* (Tamada et al. 2009; Berr et al. 2010; Guo et al. 2010; Yun et al. 2012). It has been reported previously that ATXR5 is a methyltransferase which is involved in deposition of H3K27me epigenetic marks at heterochromatin and it is then condensed eventually leading to repression of the respective gene (Jacob et al. 2009). Depending on genotyping in important recombinants of cv. Terese x *late3-1* F_2 population, the orthologous positional candidate in pea, i.e. *PsSDG15/PsATXR5* (*PsATX6* marker in this study, PsCam038017) was mapped slightly upstream of *BTB1* marker in the genetic map of *LATE3* loci on pea *LGIII* (see Figure 4.3 B).

In the present study, phylogenetic analysis of the five main classes of SET DOMAIN GROUP PROTEINS of *A. thaliana* (Springer et al. 2002) was carried out by blastP search using *A. thaliana* protein sequences against the *M. truncatula* and pea protein databases (Figure 5.1, Table 5.3) in order to clarify identity and relatedness of MtSDG10, PsSDG10, MtSDG15 and PsSDG15. The SET domain present in these methyltransferase proteins assist in various types of mono, di and tri methylation activity of H3 which are mentioned in figure 5.1. PsSDG10 (PsCam029108), the positional candidate of LATE4 belonged to a distinct clade with the group I of the *A. thaliana* SET domain proteins. The members of group I have five different domains namely EZD1 (unknown function), EZD2 (unknown function), SANT (facilitates interaction between histone tails and histone modifying enzymes), Cys-rich and SET domain (Springer et al. 2002) and results of sequence alignment have shown higher degree of conservation for these domains among the three species (appendix sequence alignment A2.1). In contrast, PsSDG15 (PsPsCam038017) which was the positional candidate of LATE3 formed clade with members of group IV SET domain proteins. The members of group IV proteins have two functional domains namely PHD (important for chromatin regulation) and SET domains (Springer et al. 2003) and sequence alignment revealed greater level of conservation for these domains between the three species as well (appendix sequence alignment A2.2).

Members of both group I and IV were highly conserved as all the three species had a single copy of these genes (Table 5.3). On the other hand, the other three groups, i.e., II, III and V showed divergence since deletion/duplication of genes for various members belonging to these three groups was observed between *A. thaliana*, *M. truncatula* and pea. For example, orthologue of AtSDG24 of group II was absent in the other two species. A similar situation

Table 5.3. SET domain group (SDG) proteins in *A. thaliana*, *M. truncatula* and *P. sativum*. *A. thaliana* accession numbers and their respective group information were taken from Springer et al. (2003), *M. truncatula* and *P. sativum* accession numbers have been collected through blastP search using *A. thaliana* protein sequences in *Medicago truncatula* Mt4.0v1 database (https://phytozome.jgi.doe.gov/pz/portal.html#) and Pea RNA seq gene atlas database (http://bios.dijon.inra.fr/FATAL/cgi/pscam.cgi) respectively.

A. thaliana name	Other Names	*A. thaliana* accession	*M. truncatula* name	*M. truncatula* accession	*P. sativum* name	*P. sativum* accession	Group
AtSDG1	CLF	At2g23380	MtSDG1	Medtr5g016870	PsSDG1	PsCam056516	I
AtSDG5	MEA	At1g02580	MtSDG5	Medtr1g086980	PsSDG5	PsCam049340	
AtSDG10	EZA1	At4g02020	MtSDG10	Medtr7g055660	PsSDG10	PsCam029108	
AtSDG4	ASHR3	At4g30860	MtSDG4	Medtr0071s0100	PsSDG4	PsCam042272	II
AtSDG7	ASHH3	At2g44150	MtSDG7	Medtr6g059310	PsSDG7	PsCam035825	
AtSDG8	ASHH2	At1g77300	MtSDG8	Medtr3g088625	PsSDG8	PsCam044812	
AtSDG24	ASHH4	At3g59960	MtSDG24	-	PsSDG24	-	
AtSDG26	ASHH1	At1g76710	MtSDG26a	Medtr1g098370	PsSDG26	PsCam045522	
			MtSDG26b	Medtr1g098000	-	-	
AtSDG14	ATX3	At3g61740	MtSDG14a	Medtr8g027725	PsSDG14a	PsCam043994	III
			MtSDG14b	Medtr7g117355	PsSDG14b	PsCam059591	
AtSDG16	ATX4	At4g27910	MtSDG16	Medtr3g091310	PsSDG16	PsCam035512	
AtSDG25	ATXR7	At5g42400	MtSDG25	Medtr2g036060	PsSDG25	PsCam006807	
AtSDG27	ATX1	At2g31650	MtSDG27	Medtr7g021365	PsSDG27	PsCam048195	
AtSDG29	ATX5	At5g53430	MtSDG29	Medtr1g008230	PsSDG29	PsCam034344	
AtSDG30	ATX2	At1g05830	MtSDG30	-	PsSDG30	-	
AtSDG15	ATXR5	At5g09790	MtSDG15	Medtr3g095840	PsSDG15	PsCam038017	IV
AtSDG34	ATXR6	At5g24330	MtSDG34	Medtr1g007670	PsSDG34	PsCam034762	
AtSDG3	SUVH2	At2g33290	MtSDG3	Medtr1g035420	PsSDG3	PsCam035414	V
AtSDG6	SUVR5	At2g23750	MtSDG6	Medtr5g018850	PsSDG6	PsCam000035	
AtSDG9	SUVH5	At2g35160	MtSDG9	Medtr5g013420	PsSDG9	PsCam048168	
AtSDG11	SUVH10	At2g05900	MtSDG11	Medtr7g088370	PsSDG11	PsCam049633	
AtSDG13	SUVR1	At1g04050	MtSDG13	Medtr7g098390	PsSDG13	PsCam035761	
AtSDG17	SUVH7	At1g17770	MtSDG17	Medtr8g042750	PsSDG17	PsCam034281	
AtSDG18	SUVR2	At5g43990	MtSDG18a	Medtr1g069570	PsSDG18	PsCam020358	
			MtSDG18b	Medtr1g069570	-	-	
AtSDG19	SUVH3	At1g73100	MtSDG19	Medtr3g011440	-	-	
AtSDG20	SUVR3	At3g03750	MtSDG20	Medtr6g027690	PsSDG20	PsCam035781	
AtSDG21	SUVH8	At2g24740	-	-	-	-	
AtSDG22	SUVH9	At4g13460	MtSDG22	Medtr8g070070	PsSDG22	PsCam048873	
AtSDG23	SUVH6	At2g22740	MtSDG23	Medtr4g010830	PsSDG23	-	
AtSDG31	SUVR4	At3g04380	MtSDG31	Medtr1g080340	PsSDG31	PsCam036130	
AtSDG32	SUVH1	At5g04940	-	-	PsSDG32	PsCam045150	
AtSDG33	KYP, SUVH4	At5g13960	MtSDG33a	Medtr7g084090	PsSDG33	PsCam042602	
			MtSDG33b	Medtr6g061270	-	-	

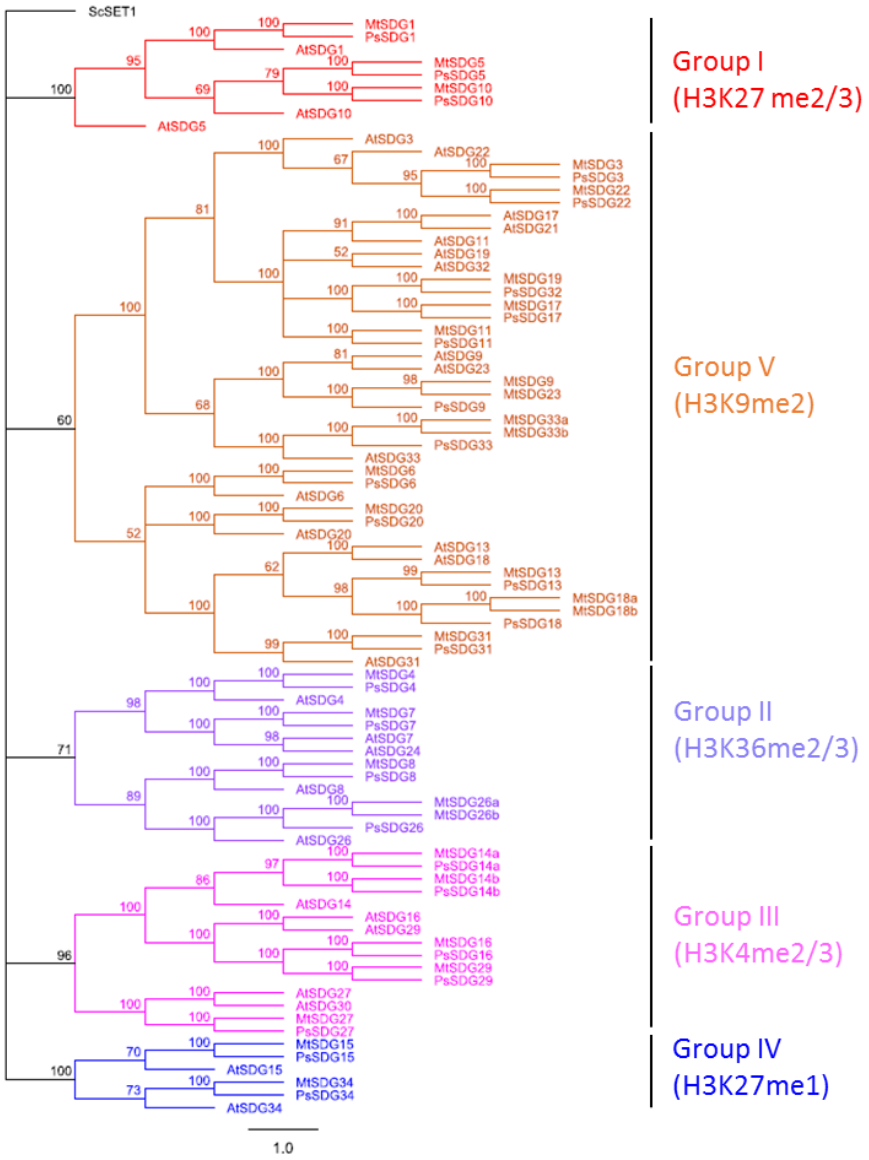


Figure 5.1. Phylogenetic tree (maximum parsimony) of *A. thaliana* SET domain group proteins and corresponding *P. sativum* and *M. truncatula* proteins. Full length protein sequences from these three species were used for construction of the phylogram. Information on protein names and accession numbers are given in table 5.1. Bootstrap values have been gained through 1000 trees and these are mentioned as a percentage next to each branch. The phylogram has been rooted using *Saccharomyces cerevisiae* SET domain protein 1 (ScSET1).

occurred for AtSDG30 of group III. Duplication was found in *M. truncatula* for MtSDG26 and MtSDG33 of group II and V respectively.

In order to look for potential functional polymorphism in *PsSDG10* and *PsSDG15*, both these genes were amplified from genomic DNA (gDNA) by using overlapping primers and then sequenced (Figure 5.2 A-B). However, sequencing results did not reveal any polymorphism

between the WT and respective *late3* and *late4* mutants. Therefore, both these genes were eliminated as potential candidates for *LATE3* and *LATE4* genes.

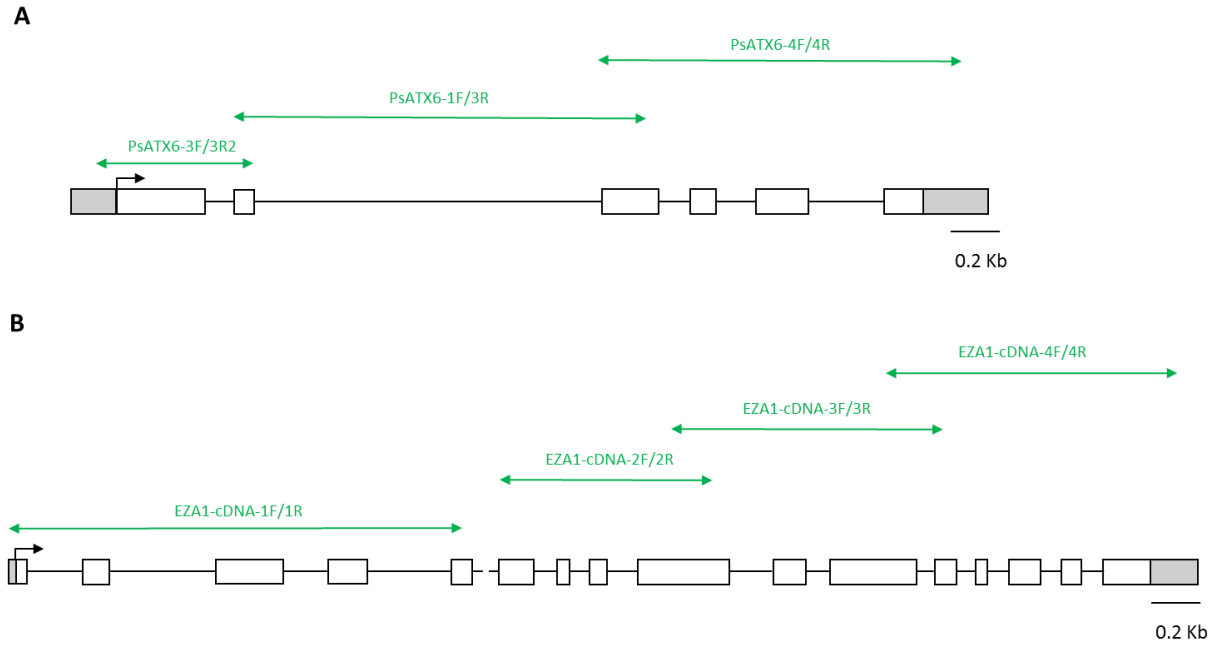


Figure 5.2. (A-B) Gene structure and position of overlapping primers used to sequence genomic DNA of the candidate gene *PsSDG15/PsATXR5* (Pscam038017) on *PsLGIII and PsSDG10/PsEZA1* (Pscam049340) on *PsLGV*. Shaded boxes represent UTR regions, white boxes represent exons, straight lines indicate introns. Arrow placed between 5'UTR and exon 1 represents translation initiation site. Primer names are mentioned in green color, doube arrow green color lines indicate amplified PCR fragments. For primer names and seqeuences, refer to the table A1.3 in appendix.

5.4.2 RNA sequencing data reveals pea Cyclin dependent kinase 8 (CDK8) as potential candidate for LATE3

Since the initial screening of different functional/positional candidate genes within the defined interval did not reveal causal mutations, therefore RNA sequencing was carried out in order to speed up this process. At first, a low coverage run was carried out using NGB5839, *late3-1* and *late4-1* genotypes (Table 5.4). All the pea transcripts within markers *BTB1* to *SPS1* in *PsLGIII* and *S62* to *OH1* in *PsLGV* as shown in table 5.1 and 5.2 were used as reference sequences. Pair end reads from the low coverage run were aligned against respective reference sequences and the resultant contigs were inspected for potential sites of polymorphism. The number of reads and percentage of reference sequence coverage obtained have been presented in appendix table A2.1 and A2.2.

Table 5.4 Total number of reads obtained for various genotypes from low and high coverage RNA sequencing runs. Data courtesy V. Hecht (unpublished).

	Number of reads	
Genotype	Low coverage run	High coverage run
WT	2196372	-
late3-1	1301420	14535740
late4-1	2035757	-
late5	-	14126094

In the low coverage run, majority of the transcripts located within defined interval in *PsLGIII* developed contigs with less than 100 reads for both WT and *late3-1* genotypes which were 43 and 50 transcripts respectively (Table A2.1). Similarly, 36 out of the 62 transcripts of WT had less than 80% coverage for the reference sequence whereas 52 of these transcripts represented the same category in *late3-1* (Table A2.1). Moreover, the number of transcripts that did not generate any contigs in WT and *late3-1* were three and four respectively. This would mean that these genes were not expressed or were below the level of detection in the selected tissue at the chosen point of time. No potential site of causal mutation was detected from these set of data.

For the defined interval of *PsLGV*, it was found that 33 and 37 transcripts from WT and *late4-1* genotypes respectively generated contigs with less than 100 reads (Table A2.2). All of these transcripts did not have more than 80% coverage as well. There were seven and four transcripts from WT and *late4-1* genotypes respectively that did not generate any contigs. This would again imply that the respective genes had no or very low expression in the selected tissue at the time point of choice. Any putative SNPs that can be regarded as causal mutation in *late4* could not be detected in this case as well.

It was apprehended from the first RNA sequencing run that the coverage of the run is a crucial factor. It was also expected that with higher coverage many of the transcripts having very low expression in the given tissue would possibly be detected. Besides, such higher coverage run is supposed to increase the number of reads even for the high and medium expressed genes. Therefore, a second RNA sequencing run using less diluted cDNA library (Hecht V, personal communication) was performed for *late3-1* and isogeneic *late5* mutant which has been used as WT. This second run markedly increased the number of reads compared to the first run

(Table 5.4). Besides, the number of transcripts in WT and *late3-1* genotype producing contigs with >400 reads increased considerably (Figure 5.3A). A similar pattern was observed in both the genotypes for the number of transcripts having >80% coverage (Figure 5.3B).

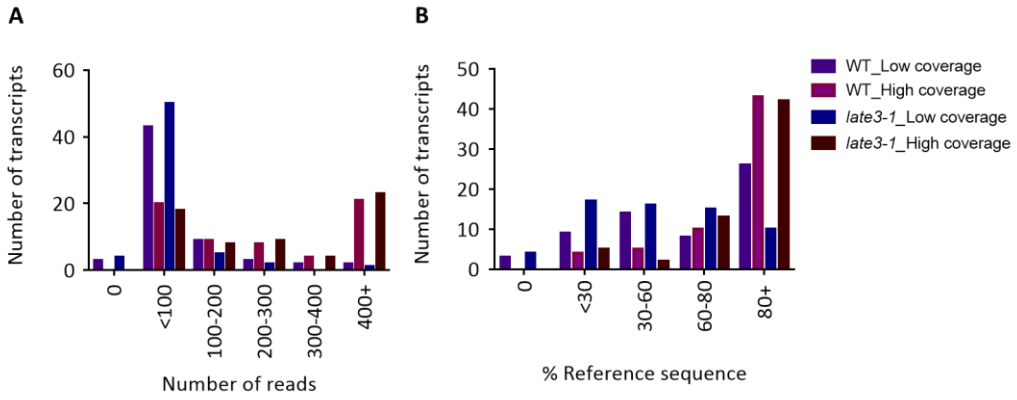


Figure 5.3. Comparison of RNA sequencing data from two different runs. **(A)** Number of transcripts showing varying range of read numbers aligned against the reference sequence in low and high coverage runs. **(B)** Number of transcripts showing varying range of percentage coverage of the reference sequence in low and high coverage runs.

One of the strong candidate gene for *late3* mutation that was considered within the defined interval of *BTB1* and *SPS1* markers in *PsLGIII* and *M. truncatula* physical map of chromosome 3 apart from *PsSDG10* was *Gibberellin 20-oxidase* (Medtr3g096500). Mutation in *GA20ox* genes was reported to show late flowering, but dwarf phenotype unlike *late3* mutants as GA is involved in stem elongation whereas overexpression of *GA20ox* promoted flowering and stem elongation (Huang et al. 1998; P et al. 1999; Schomburg et al. 2003; Mitchum et al. 2006; Rieu et al. 2008; Plackett et al. 2012). Any polymorphism was also not found in UTR/exon regions for this gene between WT and *late3-1* in the respective pea transcript, PsCam046205. Therefore, this gene was eliminated for further studies.

Another strong candidate gene was the orthologue of *M. truncatula Cyclin dependent kinase E* (*MtCDKE1*) or Cyclin dependent kinase 8 (*MtCDK8*) gene located at 44.39 Mb of chromosome 7. In *A. thaliana*, *AtCDKE1/AtCDK8* gene was found to play role in the determination of floral organ identity (Wang and Chen 2004) and *AtCDK8* mutants showed delayed flowering phenotype (Zhu et al. 2014). Besides, *CDK8* gene is a member of the highly conserved global eukaryotic transcriptional regulator called mediator complex and some

other genes of this large complex are already known to regulate flowering such as *MED8*/*12*/*13*/*18*/*25* (Allen and Taatjes 2015; Yang et al. 2016; Samanta and Thakur 2015). Upon further screening for potential causal mutations, a G to A mutation characteristic of EMS exposure was identified at 5'UTR of the pea transcript, PsCam048317 in *late3-1* genotype (Figure 5.4 A-B). A key point to mention here that the Pea RNA seq gene atlas database (http://bios.dijon.inra.fr/FATAL/cgi/pscam.cgi) currently has incorrect transcript and protein sequence for PsCam048317 as large portion of both of these sequences were found to be inserted twice, thus creating a larger sequence than actual length (Hecht V, personal

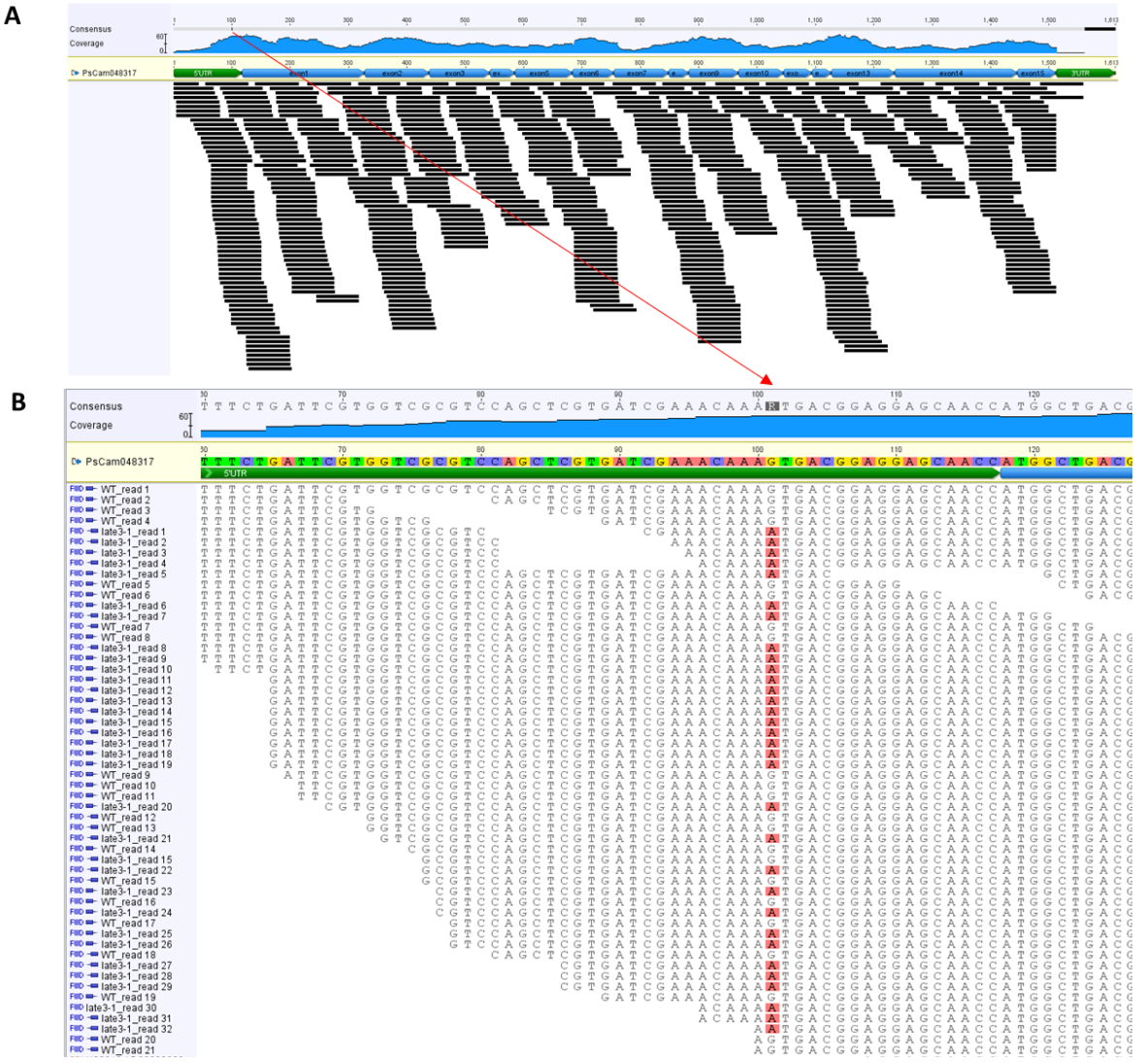


Figure 5.4. Identification of SNP at the 5'UTR of *PsCDK8* (Pscam048317) gene in *PsLGIII* through RNA sequencing. **(A)** Alignment of RNA sequencing reads from wild-type (WT) NGB6839 and *late3-1* genotype to reference sequence. **(B)** Mutation (G/A) occurred at 17bp upstream of the normal start codon of *PsCDK8* in *late3-1* genotype leading to introduction of a potential alternative start codon.

communication). From here on, any further reference of the transcript and protein sequences of *PsCDKE1*/*PsCDK8* should consider the corrected version. Since the SNP at 5'UTR of PsCam048317 was detected in this corrected version of the transcript and therefore, *PsCDKE1*/*PsCDK8* was decided to be studied further.

5.4.3 Selection of pea Cyclin C (CYCC1) as potential candidate for LATE4

It is known that the mediator complex has four different modules called head, tail, middle and CDK8. Among these, the CDK8 module is comprised of CDK8, Mediator 12 (MED12), Mediator 13 (MED13) and Cyclin C (CYCC1) which act together to repress transcription (Yang et al. 2016). The *MED12* and *MED13* have been reported to mediate various developmental processes including flowering in *A. thaliana* (Imura et al. 2012) whereas flowering phenotype of *cycc1* mutants has not been reported in any plant system yet. In light of this information and the on-going hypothesis of similar regulatory roles for *LATE3* and *LATE4* genes, the defined interval between *S62* and *OH1* markers in *PsLGV* was screened again in the syntenic physical map of *M. truncatula* even though there was no high coverage RNA sequencing data available for *late4-1* genotype. This led to the identification of *M. truncatula* gene *Cyclin C* (*MtCYCC1*, Medtr7g055650) within the aforementioned interval in *M. truncatula* chromosome 7 (Weller J and Hecht V, personal communication). The orthologous pea transcript (*PsCYCC1*) sequence accession number is PsCam050605. Depending on this result, *PsCYCC1* was chosen as the potential positional candidate for *LATE4* gene.

5.4.4. Evaluation of PsCDKE1/PsCDK8 gene as the candidate for LATE3

5.4.4.1 Phylogenetic analysis and open source expression data

Plants have evolved a wide range of cyclin dependent kinases (CDKs) which have been classified into eight different groups in *A. thaliana* (CDKA, CDKB, CDKC, CDKD, CDKE, CDKF, CDKG and CDKL) depending on their cycling binding domains (Table A2.3) (Inzé 2007; Tank and Thaker 2011). These CDKs act together with specific cyclin proteins and phosphorylate downstream target proteins which is essential to carry out various steps of cell cycle. Thus, the interactive role of CDK and cyclins assist the plants to regulate various stages of development. Among these CDKs, only CDKE1/CDK8 is known to be a component of mediator complex.

In this study, *A. thaliana* CDK8 protein sequence was used as the query for blastP search against the *M. truncatula* and pea protein databases. Phylogenetic analysis using selected proteins from sister clade and outgroups (not shown) for AtCDK8 revealed that *M. truncatula* protein, Medtr3g096960 which was located within the syntenic region in chromosome 3 as well as its orthologue in pea, PsCam048317 belonged to the same subclade with AtCDK8 (Figure 5.5). Comparison of the functional protein kinase domain of CDK8 between these three species showed high conservation (Appendix sequence alignment A2.3). These results provide indication of close evolutionary relationship for the origin and role of *CDK8* gene within these three species.

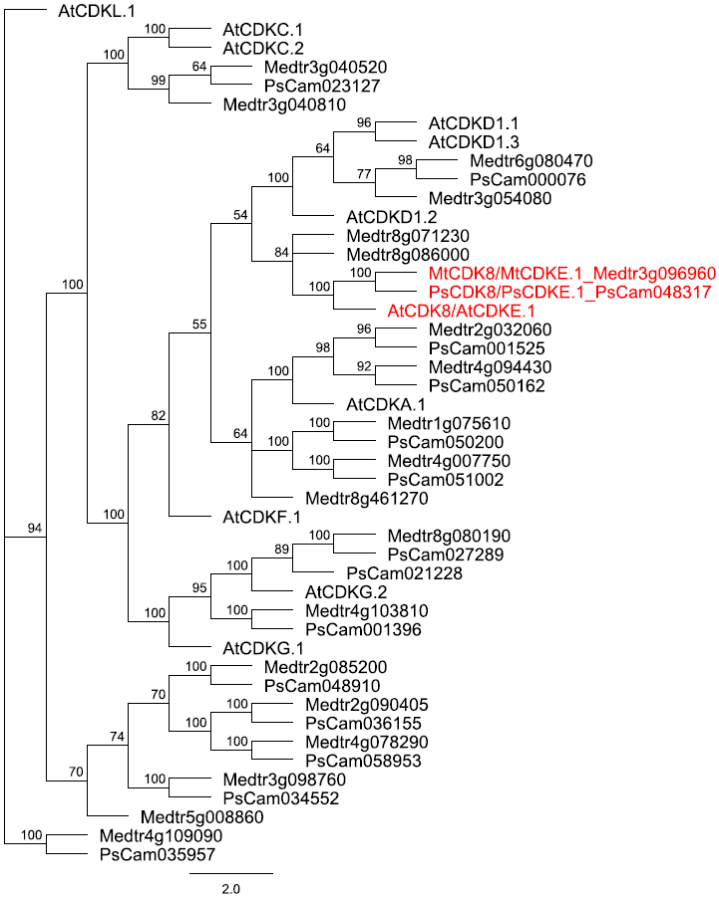


Figure 5.5. Phylogenetic tree (maximum parsimony) for PsCDK8 (accession number Pscam048317) showing that it belongs to the *A. thaliana* Cyclin dependent kinase 8 (AtCDK8)/Cyclin dependent kinase E1 (AtCDKE1) clade. Tree included selected proteins from sister clades and outgroup of *A. thaliana* CDK protein family (tree not shown, Table A2.3) and has been rooted to the AtCDKL.1 protein. Full length protein sequences for each gene were used for construction of the phylogram. Bootstrap values have been gained through 1000 trees and they are mentioned as a percentage next to each branch.

Expression atlas for the *CDK8* gene was also compiled from publicly accessible databases for *A. thaliana*, *O. sativa*, *L. japonicus*, *M. truncatula* and pea (Appendix Figure A2.3, A-E). The results from all the five species explicitly showed that this gene is expressed ubiquitously within the plant. Highest expression was observed in seed tissue in *L. Japonicus*, *M. truncatula* and pea whereas it is leaf and anther in *O. sativa* and flower organs for *A. thaliana.* Expression is also considerably high in leaf and seed of *A. thaliana* and in the seed of *O. sativa* whereas similar pattern could be seen in the flower and leaf of *L. japonicus.* Likewise, *CDK8* is expressed moderately in the leaf and floral tissues of *M. truncatula* and pea. Root and shoot tissues from all the five species also seems to have substantial amount of expression of this gene. The universal nature of expression of this gene in these five species could implicate that downregulation of this gene may have caused the modulation in various developmental and reproductive traits observed for *late3* mutants in pea (chapter 3).

Taken together, the results of phylogenetic analysis and global expression pattern of *CDK8* gene indicate that it is likely to be part of highly conserved mediator complex in pea.

5.4.4.2 Analysis of co-segregation and alternative splicing

In order to confirm the point mutation (G -> A) that was observed through RNA sequencing results at the 5'UTR of *PsCDK8* gene in *late3-1* genotype, transcript sequence of this gene in *late3-1* was generated by direct sequencing using overlapping primers (Figure 5.6 A). The WT genotype was included to verify transcript structure of pea and *M. truncatula*. Besides, the other two alleles of *late3* were added to check for presence of causal mutations. Sequencing results affirmed the putative SNP as a real event of mutation at the 5'UTR of *late3-1* which introduced a new start codon (GTG->ATG) (Figure 5.6 A, appendix figure A2.1 A). This out-of-frame mutation is likely to generate a very short ORF and the resultant hypothetical protein would be only 25 aa long which would be completely different than the PsCDK8 protein (Figure 5.6 E). The *CDK8* gene was genotyped in the important recombinants of cv. Terese x *late3-1* F_2 mapping population and it showed co-segregation with classical marker *LATE3* and as well as molecular marker *AAP2* which assured its genetic map position within the defined interval (Figure 5.6 B).

Results of cDNA sequencing from *late3-2* showed skipping of exon 13 which could potentially occur due to mutation at the acceptor splice site (Wang et al. 2015). In order to evaluate this

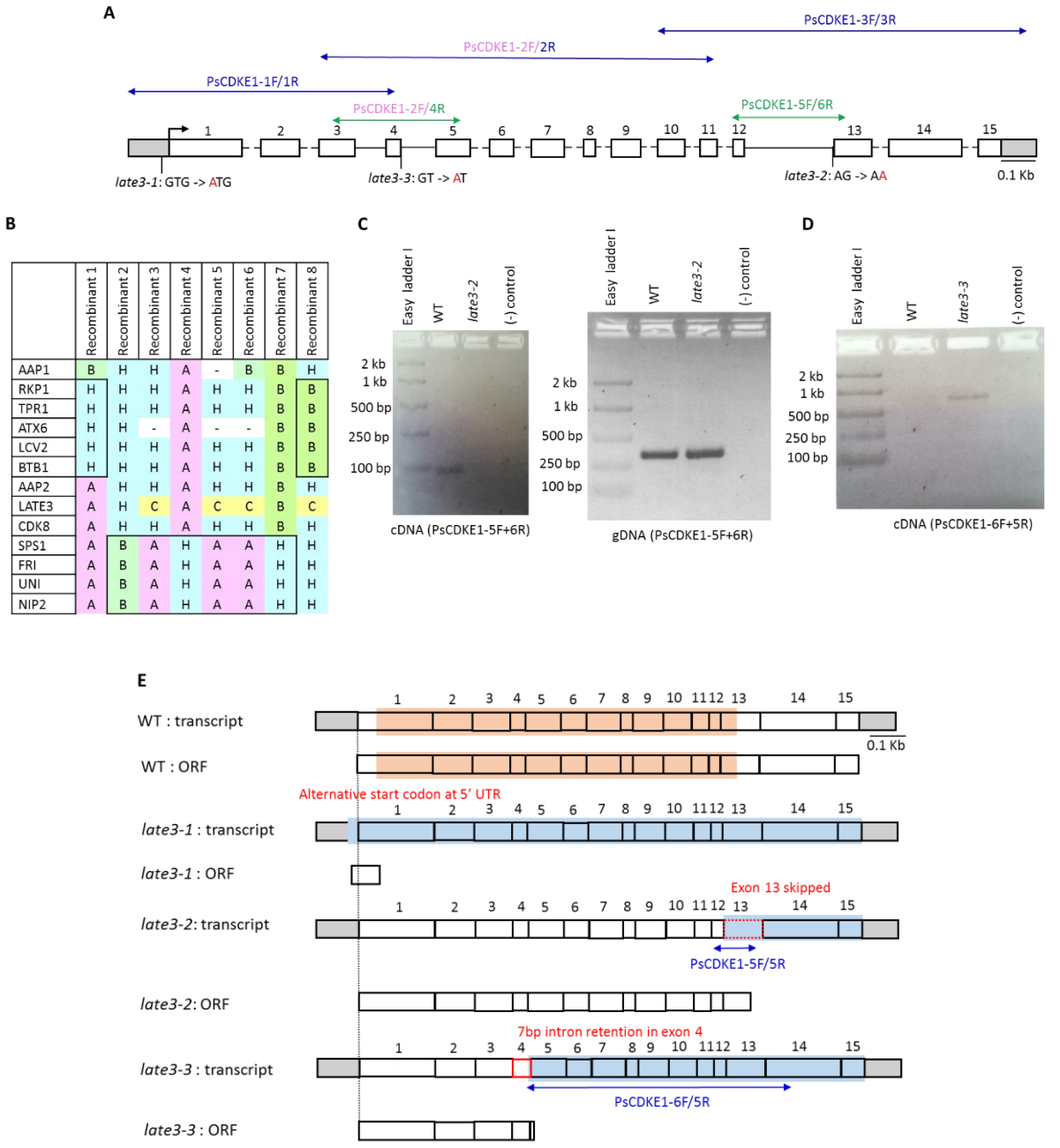


Figure 5.6. Gene structure, genotyping, alternative splicing and hypothetical ORF generation for *PsCDK8*/*PsCDKE1* (Pscam048317) gene. **(A)** Overlapping primers were designed for sequencing the *PsCDK8* gene. Gray boxes represent UTR regions, white boxes represent exons, straight lines indicate introns, Arrow placed between 5'UTR and exon 1 represent translation initiation site. Green and blue double arrow lines represent gDNA and cDNA fragments respectively. Primer names shown in green and blue colors were used for gDNA and cDNA sequencing respectively whereas those with pink color were used for both purposes. Site of mutation mentioned for *late3-1*, *late3-2* and *late3-3* alleles are located at 5'UTR, acceptor splice site of intron 12 and donor splice site of intron 4 respectively. Refer to the table A1.3 in appendix for sequences of the primer and appendix figure A2.1 for the sequences around the site of mutation. **(B)** Genotype of *CDK8* marker in the important recombinants obtained from cv. Terese x *late3-1* – F_2 population showing that it co-segregates with *LATE3* locus in *PsLGIII* and it is located within the candidate region determined between the markers *BTB1* and *SPS1*, A = Homozygous *late3-1,* B = Homozygous cv. Terese, H = Heterozygous, C = Homozygous cv. Terese / Heterozygous. **(continued next page)**

Figure 5.6 (continued) (C-D) PCR gel showing alternative splicing in *late3-2* and *late3-3* genotype. **(E)** Hypothetical open reading frame likely to be generated in *late3-1*, *late3-2* and *late3-3* mutants for the *PsCDK8* gene. Orange boxes represent major kinase domain in the WT protein. Light blue boxes represent the part of coding sequence of *late3-1*, *late3-2* and *late3-3* that are likely to be frameshifted due to mutation. Red colored dashed and plain box shown in *late3-2* and *late3-3* transcripts respectively indicate alternatively spliced exons. Primer names and double arrow lines shown in dark blue color are indicating amplified PCR fragment which were used for confirmation of alternative splicing. ORFs were predicted using NCBI ORF finder tool (https://www.ncbi.nlm.nih.gov/orffinder/).

hypothesis, primers (PsCDKE1 5F+6R) were designed on exon 12 and exon 13 to get the gDNA sequence of intron 12 from *late3-2* and WT (Figure 5.6 A). Sequencing result showed the occurrence of a G to A mutation and modulation of acceptor splice site (AG -> AA) of intron 12 which could result in generation of alternative transcript (Figure 5.6 A, A2.1 B). The same primer combination was used to amplify PCR product using cDNA from *late3-2* and WT to get further evidence on generation of splice variant. Absence and presence of PCR product in *late3-2* and WT genotypes respectively confirmed the event to alternative splicing in the *late3-2* mutant (Figure 5.6 C, A2.1 B). This mutation is likely to cause frameshift generating a truncated ORF and protein in *late3-2* than WT which would affect the major protein kinase domain of PsCDK8 as well (Figure 5.6 E).

For *late3-3*, cDNA sequencing results showed insertion of 7 bp between exon 4 and 5 which could occur due to mutation at donor splice site (Lewandowska 2013) of intron 4. To test this hypothesis, entire intron 4 was amplified using the primers PsCDKE1-2F and PsCDKE1-4R from gDNA of WT and *late3-3* genotypes and then sequenced (Figure 5.6 A). Results of this gDNA sequencing confirmed presence of a G to A mutation and thereby alteration in 5' donor splice site of intron 4 (GT->AT) in *late3-3* (Figure 5.6 A, appendix figure A2.1). Comparison of gDNA and cDNA sequences revealed the activation of a cryptic splice site (GT) located in intron 4 which resulted in retention of 7 bp of exon 4 in *late3-3* (Figure A2.1 C). In order to further prove generation of alternative transcript, forward primer (PsCDKE1-6F) was designed at 3' end of exon 4 in *late3-3* genotype which included the inserted 7 bp as well. This forward primer was used with another reverse primer (PsCDKE1-5R) to amplify the cDNA fragment within this region in WT and *late3-3* genotype (Figure 5.6 D, E). The subsequent PCR showed amplification from *late3-3* genotype, but not from WT affirming the generation of alternative transcript in the former genotype (Figure 5.6 D). Such a splice variant is likely to result in

frameshift producing a shorter ORF and protein than WT where the main protein kinase domain would also be affected (Figure 5.6 E).

Thus, analysis of relevant cDNA and gDNA sequencing as well as alternative splicing provided strong evidence on the occurrence of independent functional mutations in pea homologue of *CDK8* gene for all the three alleles of *late3* and thereby confirming this gene as *LATE3* locus. Since, all the other pea genes within the defined interval mentioned in table 5.1 did not show polymorphism in the high coverage RNA sequencing data between WT and *late3-1,* and functionally significant mutation was identified for multiple alleles of *late3-1*, therefore those genes were eliminated for further studies.

5.4.5. Evaluation of PsCYCC1 gene as the candidate for LATE4

5.4.5.1 Phylogenetic analysis and open source expression data

Cyclin proteins control various steps of eukaryotic cell cycle by regulating the activity of cyclin dependent kinases (CDK). It has been reported previously that *A. thaliana* contains 10 different classes of cyclin proteins namely Cyclin A, B, C, D, H, L, T, U, J18 and SDS (Table A2.4) (Wang et al. 2004). Among the 10 cyclin classes, five of the cyclins D, J18, L, T and SDS are plant specific. The other five classes are common between plants and animals. To date, cyclin C (CYCC1) is the only cyclin protein known to be a component of the eukaryotic mediator complex. As mentioned earlier, CYCC1 plays crucial role to negatively regulate transcription of various eukaryotic genes along with the three other members of the CDK8 module. In *A. thaliana*, there are two cyclin C proteins and they carry a cyclin N domain which is needed for their interaction with the CDK8 protein.

During the present study, *A. thaliana* cyclin C protein sequences were used for blastP search against the protein database of *M. truncatula* and pea. Phylogenetic tree built by selected sister clade and outgroups for AtCYCC1 proteins (not shown) revealed that *M. truncatula* protein, Medtr7g055650 which was found be to be located in the syntenic region of chromosome 7 (defined by *S62* and *OH1* markers in *PsLGV*) and the respective pea orthologue, PsCam050605 formed distinct clade with their *A. thaliana* orthologoues AtCYCC1.1 and AtCYCC1.2 (Figure 5.7). The key feature identified in this regard is that whereas CYCC1 duplicated in *A. thaliana*, apparently one copy of this gene originated through evolution in pea and *M. truncatula*. Comparison of major functional cyclin N domain between the three

species showed greater level of conservation (appendix sequence alignment A2.4). These results also suggested that PsCYCC1 is the component of the pea CDK8 module that is likely needed for interaction with PsCDK8.

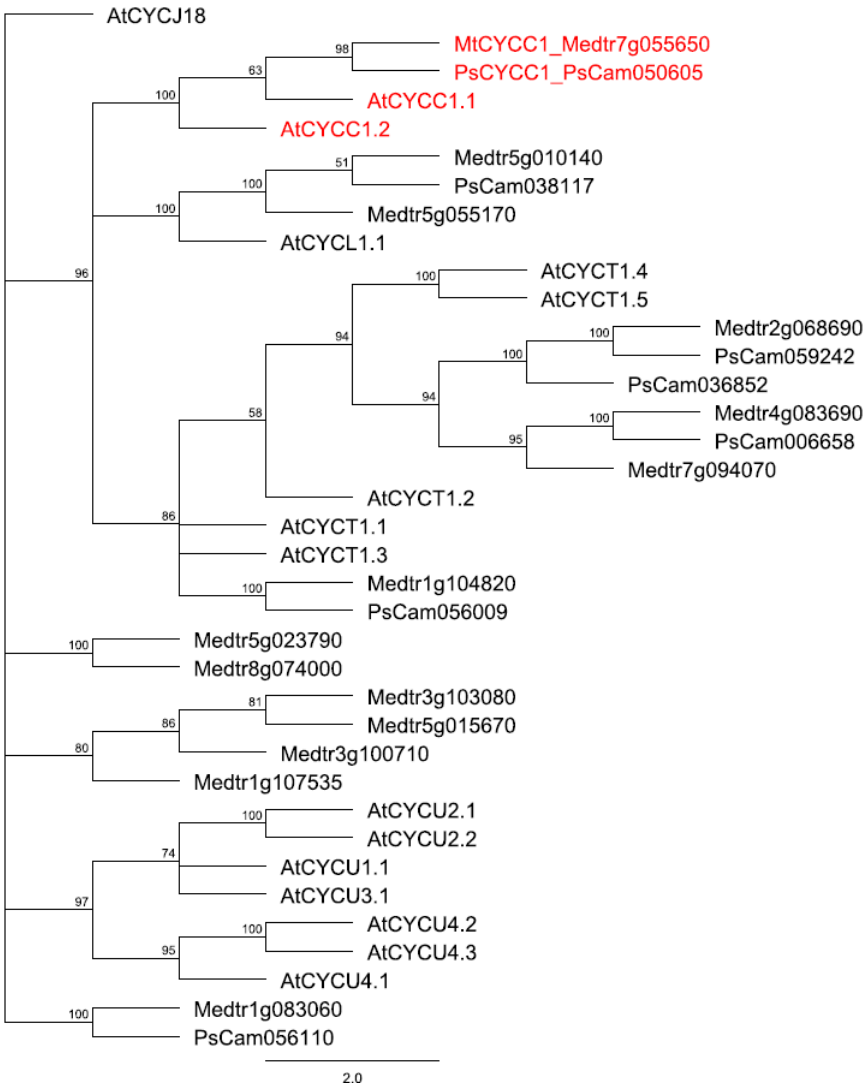


Figure 5.7. Phylogenetic tree (maximum parsimony) for PsCYCC1 (accession number Pscam050605) showing that it belongs to the same clade with *A. thaliana* Cyclin C (AtCYCC1). Tree included selected proteins from sister clades and outgroup of *A. thaliana* Cyclin protein family (tree not shown, Table A2.5) and has been rooted to the AtCYCJ18 protein. Full length protein sequences for each gene were used for construction of the phylogram. Bootstrap values have been gained through 1000 trees and they are mentioned as a percentage next to each branch.

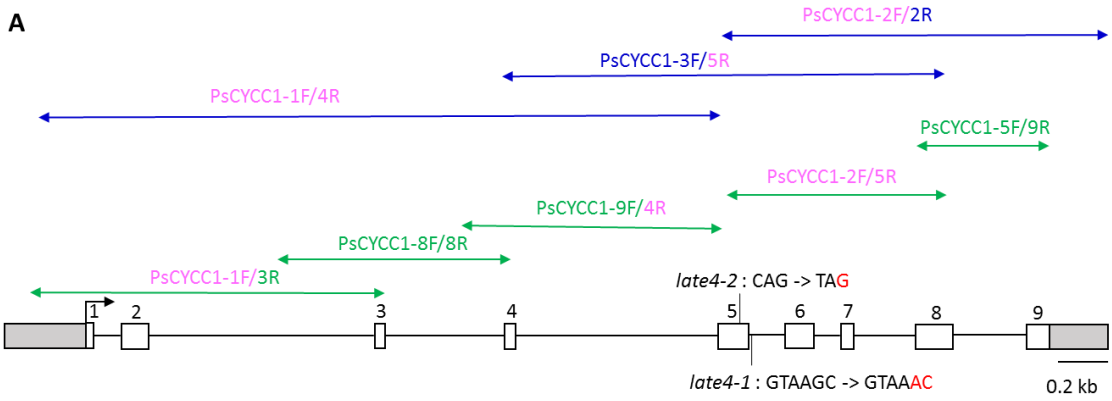
Next, gene expression data of *CYCC1* from publicly accessible databases were harvested for *A. thaliana*, *O. sativa*, *L. japonicus*, *M. truncatula* and pea (appendix figure A2.4 A-E). In *A. thaliana* and *O. sativa*, leaf and flower tissues/anther have very high expression while seed and floral tissues in *L. japonicus* show similar pattern of expression. Considerably higher

expression of the gene was observed in root and leaves of *M. truncatula* while leaves and flower are the two tissues where this gene is expressed in highest amount in pea. Root and shoot tissues also showed fair amount of expression for this gene in all the five species. In general, the expression of *CYCC1* throughout these tissues is higher than *CDK8* specially in *M. truncatula* and pea. The ubiquitous pattern of expression of *CYCC1* gene and specifically in flowering study related organs such as leaf, flower and seed are consistent with potential misregulation of the gene in *late4* mutants which probably caused the observed abnormal phenotypes across various growth stages.

5.4.5.2 Analysis of co-segregation and alternative splicing

Due to similarities in the phenotype of the *late3* and *late4* mutants and identification of *LATE3* as *PsCDK8*, it was needed to confirm that pea *Cyclin C* (*PsCYCC1*) gene as *LATE4*. To this end, overlapping primers were designed in an attempt to initially amplify the entire cDNA sequence of this gene from WT, *late4-1* and *late4-2* genotypes (Figure 5.8 A). This PCR resulted in amplification of more than one band (primer combinations PsCYCC1-3F+5R) around exon 5 in both the mutants relative to WT (Figure 5.8 A, 5.9 A). No polymorphic site was detected in both the mutants in the other regions of cDNA which yielded single bands.

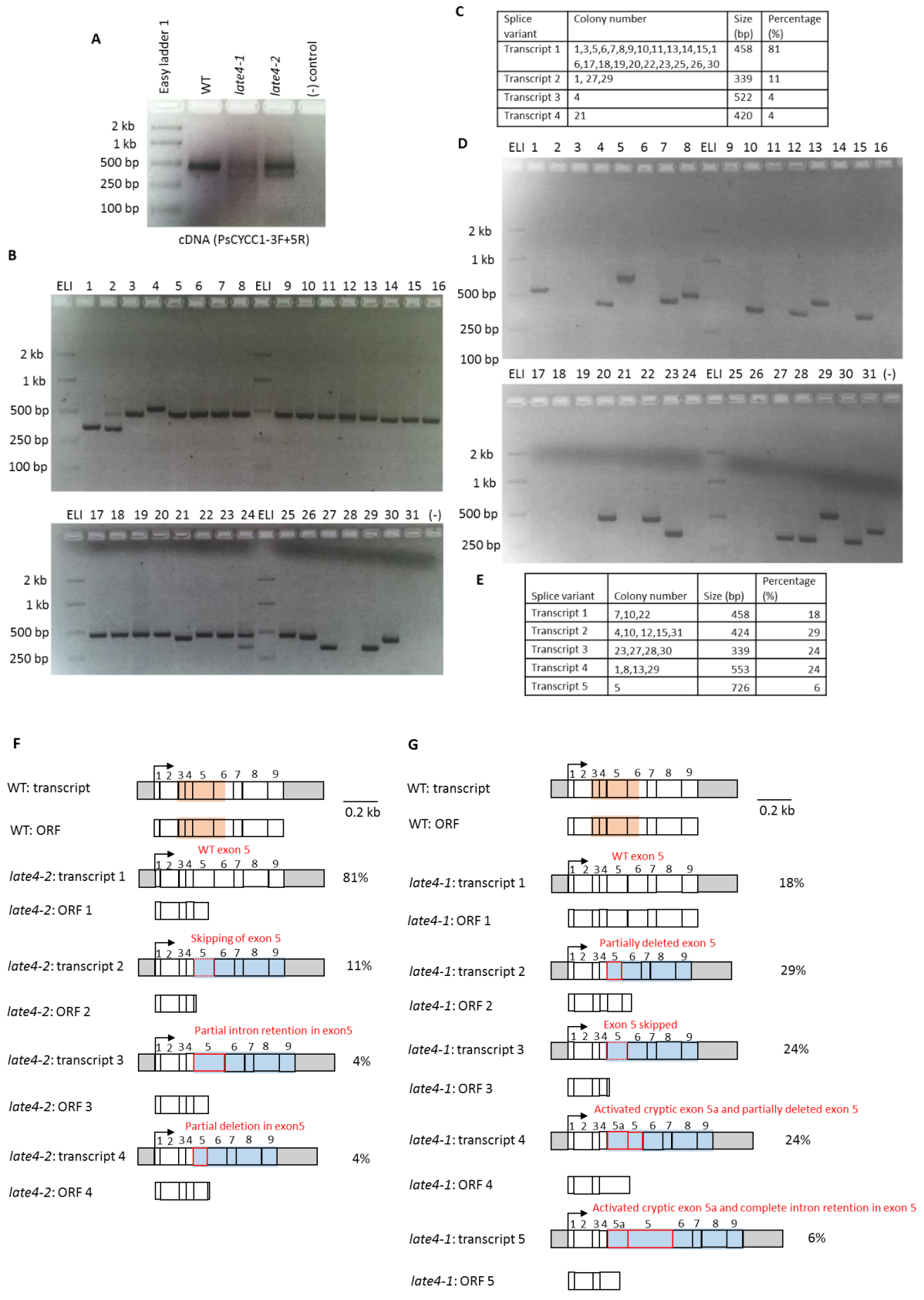
Therefore, both the alleles were sequenced again using overlapping primers on gDNA (Figure 5.8 A, A2.2 A). This sequencing result revealed presence of a C to T mutation in exon 5 of *late4-2* which introduced a premature stop codon (CAG->TAG). In contrast, a G to A mutation at +5 bp position of intron 5 was found in *late4-1* genotype which caused modification of 5' splice motif (GTAAGC->GTAAAC) (Figure 5.8 A, appendix figure A2.2 A). In *S. cerevisiae*, it was reported previously that this 5th position is invariant and mutation at this site could alter post-transcriptional modification of pre-mRNA (Lesser and Guthrie 1993). In the current study, nearly 58% of the 50 selected introns in pea also showed the presence of G at this position which was much higher than other three bases (T = 20%, A = 14%, C = 8%), thus indicating higher degree of conservation of this position across the introns in eukaryotes. Besides, results of gDNA sequencing showed that all the introns of *PsCYCC1* gene in the NGB5839 (WT) possess G in the 5th position of the 5' splice motif which gives hint that a certain type of splicing machinery might be carrying out splicing for all the introns of this gene in WT genotype. Therefore, it was assumed that G in 5th position of 5' splice motif is crucial for proper splicing of introns of *PsCYCC1* gene.



B

	Recombinant 1	Recombinant 2	Recombinant 3	Recombinant 4	Recombinant 5	Recombinant 6	Recombinant 7	Recombinant 8
ABI3	B	B	A	H	A	H	A	B
BZIP1	B	B	A	H	A	H	A	-
PAP1	B	B	A	H	A	H	A	B
GP1	B	B	A	H	A	H	A	B
OH1	B	B	A	H	A	H	A	B
DSP1	B	B	-	B	A	H	H	B
EZA1	B	B	A	B	A	H	H	B
NTF2	B	B	A	B	A	H	H	B
LATE4	C	B	A	B	A	H	H	B
CYCC1	B	B	-	-	A	H	H	B
ATPB1	B	B	A	B	A	H	H	B
ELA1	H	B	A	B	A	H	H	B
L32	H	B	H	B	H	B	H	B
S62	H	B	H	B	H	B	H	B
MCO1	H	H	H	B	H	B	H	H
ICE1	H	H	H	B	H	B	H	H
THP1	A	H	H	B	H	B	H	H

Figure 5.8. Gene structure and genotyping for *PsCYCC1* (Pscam050605). **(A)** Overlapping primers were designed for sequencing the *PsCYCC1* gene. Shaded boxes represent UTR regions, white boxes represent exons, straight lines indicate introns, arrow placed between 5'UTR and exon 1 represent translation initiation site. Green and blue double arrow lines represent gDNA and cDNA fragments respectively. Primer names shown in green and blue colors were used for gDNA and cDNA sequencing respectively whereas those with pink color were used for both purposes. Site of mutation mentioned for *late4-1* and *late4-2* alleles are located at 5th position (5' end) of intron 5 and exon 4 (introducing new stop codon) respectively. Refer to the table A1.3 in appendix for sequences of the primer and appendix figure A2.2 for the sequences around the site of mutation. **(B)** Genotype of *CYCC1* marker in the important recombinants obtained from *late4-1 x* cv. Terese – F_2 population showing that it co-segregates with *LATE4* locus in *PsLGV* and it is located within the candidate region determined between the markers *OH1* and *S62*, A = Homozygous *late4-1,* B = Homozygous cv. Terese, H = Heterozygous, C = Homozygous cv. Terese / Heterozygous.



Splice variant	Colony number	Size (bp)	Percentage (%)
Transcript 1	1,3,5,6,7,8,9,10,11,13,14,15,16,17,18,19,20,22,23,25, 26, 30	458	81
Transcript 2	1, 27,29	339	11
Transcript 3	4	522	4
Transcript 4	21	420	4

Splice variant	Colony number	Size (bp)	Percentage (%)
Transcript 1	7,10,22	458	18
Transcript 2	4,10, 12,15,31	424	29
Transcript 3	23,27,28,30	339	24
Transcript 4	1,8,13,29	553	24
Transcript 5	5	726	6

Figure 5.9. Alternative splicing and hypothetical ORF generation for *PsCYCC1* (Pscam050605). **(continued next page)**

Figure 5.9. (continued) (A) Confirmation of alternative splicing in *late4-1* and *late4-2* alleles. PCR product amplified from cDNA of *late4-1* and *late4-2* genotypes (showing multiple bands) using forward primer PsCYCC1-3F and reverse primer PsCYCC1-5R (see Figure 5.8 A for primer location) were selected for cloning whereas WT PCR product was sequenced (458 bp) directly. Refer to the table A1.3 in appendix for sequences of the primers. **(B-C)** Four different splice variants were identified from *late4-2* genotype, percentage of transcripts have been rounded up to two digits. **(D-E)** Five different splice variants were identified from *late4-1* genotype, percentage of transcripts have been rounded up to two digits. **(F-G)** Structure of hypothetical open reading frame (ORF) likely to be generated from various splice variants in *late4-2* and *late4-1* mutants for the *PsCYCC1* gene. Gray boxes for transcripts represent UTR regions, white boxes represent normal exons, red colored dashed or plain boxes represent alternatively spliced exons and orange boxes represent major cyclin N domain. Light blue boxes represent the part of coding sequence of various transcripts of *late4-1* and *late4-2* that are likely to be frameshifted due to mutation. The ORFs were predicted using NCBI ORF finder tool (https://www.ncbi.nlm.nih.gov/orffinder/). Refer to appendix figure A2.2 for the results of direct sequencing for these splice variants.

The *PsCYCC1* gene was then genotyped in six out of eight of the important recombinants as DNA/leaf samples were unavailable for two other recombinants from *late4-1* x cv Terese F_2 population. Results of the genotyping showed that *PsCYCC1* co-segregates with the classical marker *LATE4* as well as with four other molecular marker namely *DSP1*, *EZA1*, *NTF2* and *ATPB1* (Figure 5.8 B). Thus, the genotype data established the genetic position of *PsCYCC1* within the defined interval.

Since the site of mutation for *late4-1* was found in 5th position of 5' splice motif of intron 5 and multiple bands were obtained from PCR on cDNA around exon 5 (primer combinations PsCYCC1-3F+5R), therefore these bands were hypothesized to be splice variants. Similar prediction was made for the observed multiple bands of *late4-2* around the same location. To verify this idea, these PCR products from both the genotypes were cloned in *E. coli* and 31 colonies were selected from each mutants for colony PCR. Results of the colony PCR showed presence of four different types of splice variants namely transcript type 1-4 of varying lengths in *late4-2* (Figure 5.9 B-C) while there were five such splice variants namely transcript 1-5 in *late4-1* (Figure 5.9 D-E). Sequencing results from each of these transcripts showed different patterns of alternative splicing in both the *late4* mutant alleles.

WT PCR fragment from the aforementioned primer combinations had a length of 458 bp. Various patterns of alternative splicing in *late4-2* genotype is shown in Figure 5.9 F and A2.2 B-D. Here, 81% of the splice variants namely transcript type 1 were 458 bp long similar to WT.

Next, 11% of the splice variants called transcript type 2 had the entire exon 5 (119 bp) skipped generating a PCR product of 339 bp long. Transcript type 3 consisted of 4% of the splice variants having a length of 520 bp where partial retention of intron 5 occurred in exon 5. This intron was spliced out at +64 bp of intron 5 as a cryptic 5' splice motif at this location having G in 5th position (GTGAGT) was activated. Finally, another 4% of the splice variants known as transcript type 4 were 420 bp long in which case partial deletion occurred within exon 5 just 1 bp upstream of the actual site of mutation, i.e., CAG->TAG. This site of EMS mutation was found to be embedded within a cryptic 5'splice motif (GTAGGC) in *late4-2*, thus the part of the exon 5 carrying the newly introduced premature stop codon (PTC) was spliced out. These results suggest additional effect of the non sense mutation in exon 5 of *late4-2* which led to the generation of type 2, 3 and 4 transcripts. Any hpothetical proteins that may generate from all of the four potential ORFs are likely to be truncated which would eventually affect the Cyclin N domain of the PsCYCC1 protein. Further dicussion on this particular aspect would be carried out in section 5.5.1.2.

Different types of alternative splicing in *late4-1* genotype have been presented in Figure 5.9 G and A2.2 E-H. In this case, 18% of the splice variants named as transcript type 1 were similar in length to WT (458 bp). 29% of the splice variants called transcript type 2 were 424 bp long due to 34 bp partial deletion of exon 5 which occurred because of the activation of a cryptic 5'splice motif located within exon 5 having G in 5th position (GCTAGG). Next, 24% of the splice variants known as transcript type 3 were found to have entire exon 5 (119 bp) skipped resulting in a 339 bp product. Transcript type 4 was comprised of 24% of the splice variants having a length of 553 bp and they were found to possess a 129 bp long cryptic exon namely exon 5a which was embedded within 810 bp long intron 4 (from +117 bp to +290 bp of intron 4) and it was bordered by a cryptic 3' splice site (AG) and 5' splice motif having G in 5th position (GTAAGA). The transcript type 4 also had 34 bp deletion of exon 5 similar to the situation mentioned eariler for transcript type 2. Lastly, 6% of the alternatively spliced transcripts named transcript type 5 were 726 bp long where presence of cryptic exon 5a (129 bp) occurred similar to the scenario of transcript type 4. In addition, these type of splice variants also retained the entire 139 bp of intron 5 in exon 5. Any hypothetical proteins generated by the subsequent ORFs apart from transcript type 1 likely to be truncated damaging the major Cyclin N domain of PsCYCC1 protein. In other words, certain proportion of WT PsCYCC1

protein seem to be expressed in *late4-1* due to the presence of transcript type 1. This would be discussed further in section 5.5.1.2.

Thus, analysis of colony PCR results provided strong evidence of independent causal mutations in pea homologue of *CYCC1* gene for both alleles of *late4* mutation which established this gene as *LATE4* locus.

5.4.6. Mediator complex subunits in M. truncatula and pea

Different subunits of the mediator complex in *A. thaliana* have so far been reported to play role in different biological processes including flowering (Kidd et al. 2011; Samanta and Thakur 2015; Malik et al. 2017). Based on this understanding about the importance of mediator complex in plants, various subunits of *M. truncatula* and pea mediator complex have been identified in this study (Table 5.4) by reciprocal blastP search using previously reported *A. thaliana* mediator complex proteins (Figure 5.10) (Backstrom et al. 2007; Mathur et al. 2011; Malik et al. 2017).

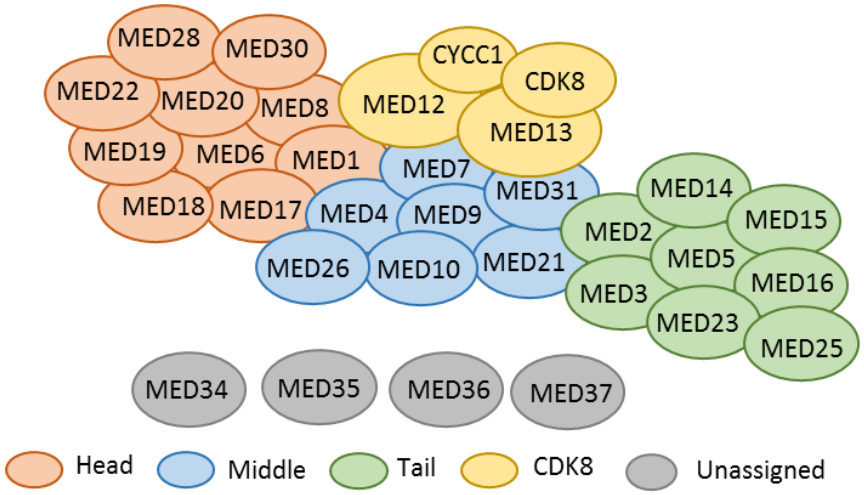


Figure 5.10. Schematic representation of various modules of *A. thaliana* mediator complex. MED34, MED35, MED36 and MED37 are plant specific subunits. Figure adapted from Malik et al., 2017.

Results of this study revealed presence of supplementary copy (-ies) of various subunits in one of the three species (Table 5.4). For example, MED22 had an additional copy in *A. thaliana* relative to the two other species. Similar situation was observed for CYCC1 and MED26. In contrast, extra copies was found to be present in *M. truncatula* for subunits MED4, MED11,

Table 5.4. Mediator complex subunits in *A. thaliana*, *M. truncatula* and *P. sativum*. *A. thaliana* accession numbers and their respective group information were taken from Mathur et al. 2011, *M. truncatula* and *P. sativum* accession numbers have been collected through blastP search in *M. truncatula* Mt4.0v1 database and Pea RNA seq gene atlas database respectively.

Present in all genotypes
Present in *A. thaliana* only
Present in *M. truncatula* only
Present in *P. sativum* only
Absent in *A. thaliana*, present in *M. truncatula* and *P. sativum*
Absent in *M. truncatula*, present in *A. thaliana* and *P. sativum*
Absent in *P. sativum*, present in *A. thaliana* and *M. trunctula*

Module name	Subunit name	*Arabidopsis thaliana*	*Medicago truncatula*	*Pisum sativum*
Head	MED6	AT3G21350	Medtr8g015230	PsCam039240
	MED8	AT2G03070	Medtr5g013930	PsCam050028
	MED11-1	AT3G01435	Medtr1g023330	PsCam043799
	MED11-2	-	Medtr8g106470	-
	MED17	AT5G20170	Medtr4g114100	PsCam048221
	MED18-1	AT2G22370	Medtr2g438770	PsCam051584
	MED18-2	-	Medtr5g044730	-
	MED19-1	AT5G12230	Medtr8g028015	PsCam000282
	MED19-2	AT5G19480	Medtr8g079890	PsCam001661
	MED20-1	AT3G28230	Medtr2g063927	PsCam052097
	MED20-2	AT4G09070	Medtr4g125510	PsCam056338
	MED22-1	AT1G16430	Medtr2g062860	PsCam045952
	MED22-2	AT1G07950	-	-
	MED28	AT3G52860	Medtr8g092180	PsCam046291
	MED30	AT5G63480	Medtr3g099130	PsCam051144
Kinase	CYCC1-1	AT5G48640	Medtr7g055650	Pscam050605
	CYCC1-2	AT5G48630	-	-
	CDK8	AT5G63610	Medtr3g096960	Pscam048317
	MED12-1	AT4G00450	Medtr7g077700	PsCam020904
	MED12-2	-	Medtr8g030800	PsCam013006
	MED13	AT1G55325	Medtr3g083500	PsCam042334
Middle	MED4-1	AT5G02850	Medtr8g075790	PsCam010753
	MED4-2	-	Medtr6g089780	-
	MED7-1	AT5G03220	Medtr6g471120	PsCam023350
	MED7-2	AT5G03500	Medtr2g094720	-
	MED9	AT1G55080	Medtr8g017180	PsCam051501
	MED10-1	AT1G26665	Medtr8g085570	PsCam007448
	MED10-2	AT5G41910	-	PsCam013322
	MED10-3	-	-	PsCam001908
	MED10-4	-	-	PsCam000929
	MED21-1	AT4G04780	Medtr7g037020	-
	MED21-2	-	Medtr1g107580	-
	MED26-1	AT3G10820	Medtr7g089140	PsCam036814
	MED26-2	AT5G05140	Medtr3g095380	PsCam049606
	MED26-3	AT5G09850	-	-
	MED31	AT5G19910	Medtr1g102540	PsCam027544

(continued next page)

Table 5.4 (continued)

Module name	Subunit name	*Arabidopsis thaliana*	*Medicago truncatula*	*Pisum sativum*
Tail	MED2/29/32	AT1G11760	Medtr3g116260	PsCam039978
	MED3/27	AT3G09180	Medtr7g106200	PsCam030836
	MED5-1/MED24-1/MED33-1	AT3G23590	Medtr2g021400	PsCam020903
	MED5-2/MED24-2/MED33-2	AT2G48110	Medtr2g105810	PsCam000396
	MED5-3/MED24-3/MED33-3	-	Medtr7g094550	PsCam027135
	MED5-4/MED24-4/MED33-4	-	Medtr7g074290	PsCam006844
	MED14-1	AT3G04740	Medtr7g115190	PsCam044823
	MED14-2	-	Medtr1g087060	-
	MED15-1	AT1G15780	Medtr2g104400	PsCam042386
	MED15-2	AT1G15770	Medtr2g104430	PsCam013021
	MED15-3	-	Medtr2g104450	PsCam043962
	MED15-4	-	Medtr4g009740	PsCam009666
	MED15-5	-	-	PsCam054856
	MED15-6	-	-	PsCam053663
	MED16	AT4G04920	Medtr4g088965	PsCam016884
	MED23	AT1G23230	Medtr1g031910	PsCam042359
	MED25	AT1G25540	Medtr5g054680	PsCam005583
Unassigned	MED34	AT1G31360	Medtr5g026590	PsCam044095
	MED35-1	AT1G44910	Medtr1g017520	PsCam042441
	MED35-2	AT3G19670	Medtr1g106025	PsCam002035
	MED35-3	AT3G19840	Medtr4g074440	PsCam048831
	MED36-1	AT4G25630	Medtr1g114000	PsCam009707
	MED36-2	AT5G52470	Medtr6g047580	PsCam001477
	MED36-3	-	Medtr8g030490	PsCam009720
	MED36-4	-	Medtr6g047800	PsCam009786
	MED36-5	-	-	PsCam010791
	MED37-1	AT1G09080	Medtr8g099945	PsCam045359
	MED37-2	AT5G28540	Medtr8g099795	PsCam042606
	MED37-3	AT5G02490	Medtr7g099680	PsCam044083
	MED37-4	AT5G02500	Medtr7g024390	PsCam036608
	MED37-5	AT5G28540	Medtr7g024580	PsCam036503
	MED37-6	AT5G42020	Medtr4g063720	PsCam014100

*MED37 has many members in *M. truncatula* and *P. sativum*, for simplicity only 6 are mentioned here

MED14, MED18 and MED21. Likewise, supplementary copies of MED10, MED15 and MED36 subunits were identified in pea.

It was also observed that additional copy (-ies) of certain subunits were absent in one out of the three species. For instance, extra copies of MED5, MED12, MED15, MED36 and MED37 subunits were present in *M. trunctula* and pea compared to *A. thaliana*. A similar result was obtained for MED10 in *A. thaliana* and pea in comparison to *M. truncatula*. For MED7 and MED21, supplementary copies were identified in *A. thaliana* and *M. truncatula* relative to pea. All other subunits of the mediator complex were found to be conserved across the three species.

Taken together, these results suggest origin of both conserved and species specific mediator complex subunits in the three species.

5.4.7 Phenotypic characterization of mutants for CYCC1 gene reveals likely conservation of function of this gene in A. thaliana

In *A. thaliana*, it was previously found that mutation in three of the four genes of CDK8 module, i.e., *CDK8*, *MED12* and *MED13* results in late flowering phenotype (Wang and Chen 2004; Imura et al. 2012; Zhu et al. 2014). Information on flowering phenotype in *AtCYCC1* gene has not been reported so far. However, in the present study, forward genetics studies of the late flowering *late4* mutants led to the identification of *PsCYCC1* as the underlying gene in pea. Since *AtCYCC1* is part of the CDK8 module in *A. thaliana*, so mutation in this gene was hypothesized to cause delay in flowering similar to the other three genes. Therefore, it was decided to study the phenotypes of known T-DNA homozygous mutant lines of *CYCC1* gene in *A. thaliana.* Such studies were also expected to reveal whether there are any similarities in the phenotype of respective mutants of this gene between the *A. thaliana* and pea. To this end, seeds for two homozygous T-DNA insertional mutants (in Col-0 background) of *AtCYCC1* gene namely *cycc1-1* and *cycc1-2* were ordered and grown.

Among these two lines, *cycc1-1* (SALK_053291C) is a homozygous knockout line for *AtCYCC1.1* gene (AT5G48640) with T-DNA insertion located at 5'UTR of the gene. In contrast, *cycc1-2* (SALK_039400C) is a knockout line for both *AtCYCC1.1* (AT5G48640) and *AtCYCC1.2* (AT5G48630) which has T-DNA insertion in the small intergenic region between the two genes (Zhu et al. 2014). Initially, the putative *cycc1-1* and *cycc1-2* mutants and WT were genotyped to confirm that the mutants are homozygous and have the insertion as desired. Results of the genotyping using gene specific primers showed the presence of PCR bands only in the WT (Col-0), but not in the *cycc-1* and *cycc1-2* mutants whereas the opposite results were obtained for insertion specific primers (Figure 5.11 A-D). Hence, these results suggest that the respective lines carried the expected T-DNA insertion and it is worth to use these homozygous lines further for characterization of phenotypes.

Both the *cycc1-1* and *cycc1-2* mutants of *A. thaliana* flowered significantly late compared to their WT counterpart (Figure 5.12 A-B, $p < 0.05$) similar to what was observed for *late4-1* and *late4-2* in pea (chapter 3). These mutants in *A. thaliana* were also significantly taller than the

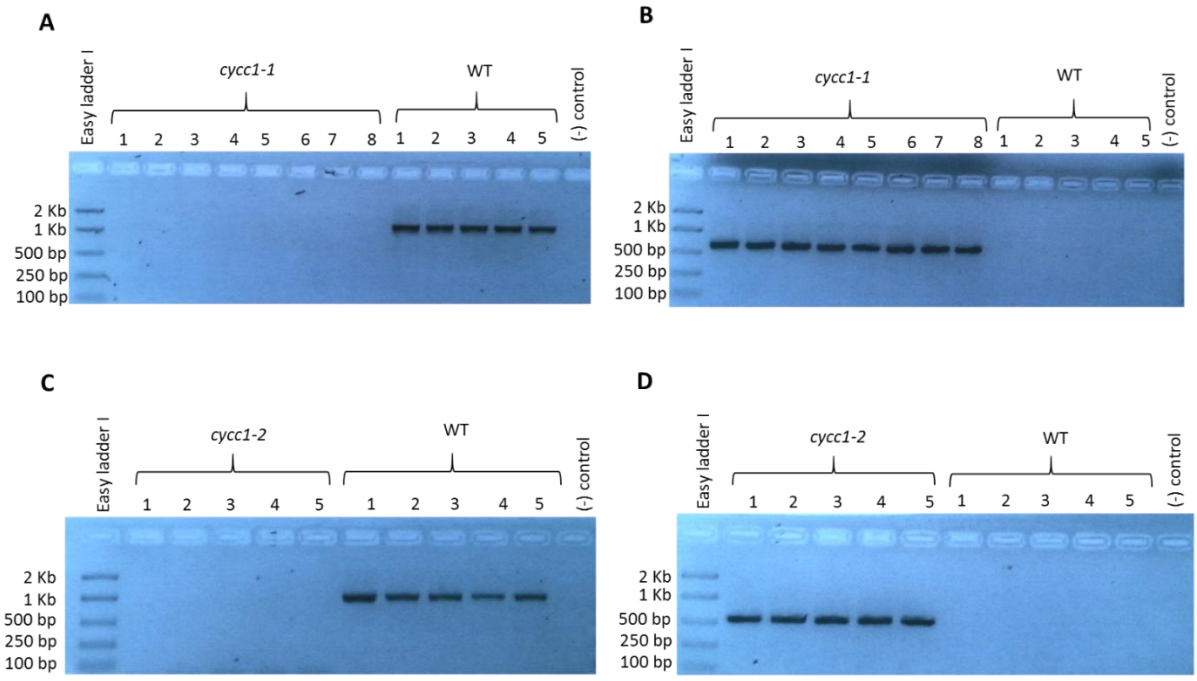


Figure 5.11. Genotyping of *A. thaliana* T-DNA insertion lines for *ATCYCC1* gene. **(A)** Gene specific and **(B)** insertion specific PCR gel for T-DNA line *cycc1-1* (SALK-039400). **(C)** Gene specific and **(D)** insertion specific agarose gel for T-DNA line *cycc1-2* (SALK_053291). Columbia-0 was used as the wild type (WT). Primers for this experiment are mentioned in table A1.4.

WT and they possessed significantly greater aerial branching (Figure 5.12 A, C, D; $p < 0.05$) as was the case for the two mutant alleles of *late4* in pea. In addition, it was observed that the siliques appeared along the main stem at an angle (~60°) in WT while more or less perpendicular appearance for the same organs in the mutants was found (Figure 5.12 A).

Modulation in the morphology of the siliques were also quantified at various growth stages in these *A. thaliana* mutants. For instance, the first silique was significantly shorter in the *cycc1-1* and *cycc1-2* mutants compared to WT while the pedicle of this silique showed opposite feature (Figure 5.13 A-C, $p < 0.05$). These siliques from the mutants did not generate any seeds. To compare phenotype of this reproductive structure at a later stage, the 10th silique was studied as a representative of that stage. In this case, mutant exhibited significantly greater length for both the silique and pedicle relative to WT (Figure 5.13 D-F, $p < 0.05$); however the siliques were much thinner in the mutants than the WT. The number of siliques along the main stem were significantly lower in the mutants than the WT (Figure 5.13 E, $p < 0.05$). Seeds of both the *cycc1* mutants of *A. thaliana* were also smaller in comparison to WT as was the case for the pea *late4* mutants.

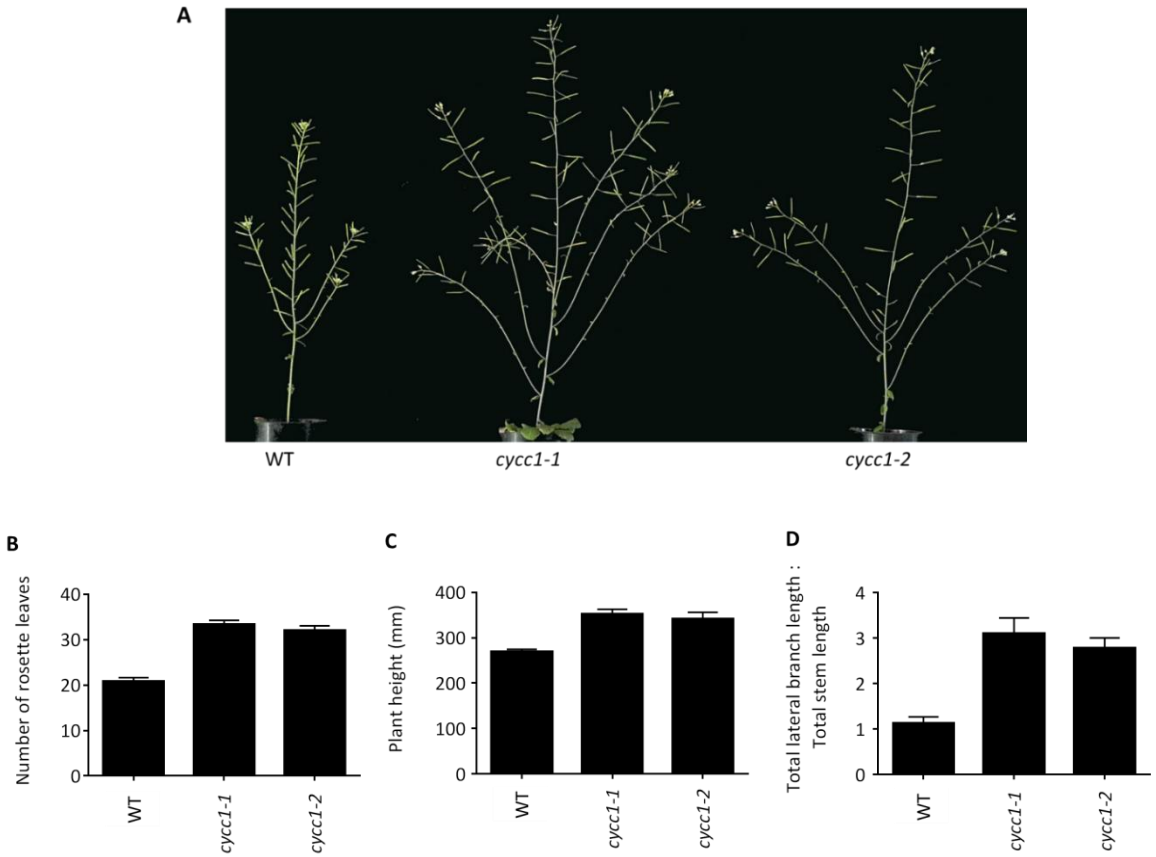


Figure 5.12. Modulation in various traits between the *A. thaliana* WT Columbia-0 and T-DNA line *cycc1-1* and *cycc1-2*. **(A)** Representative 74 day old WT, *cycc1-1* and *cycc1-2* mutants grown under LD. **(B)** Initiation of flowering as a count of the number of rosette leaves, data represents mean ± SE for n = 5-8 plants. **(C)** Plant height measured at the time of harvest 90 days after sowing, data represents mean ± SE for n = 5-8 plants. **(D)** Total branching in the WT and mutants presented as a ratio of total lateral branch length and total stem length, total length of all lateral branches such as primary, secondary branches and total stem length were measured at the time of harvest, i.e., 90 days after sowing, data represents mean ± SE for n = 5-8 plants.

The delayed flowering phenotype of *cycc1* mutants of *A. thaliana* observed in the present study was consistent with that of already reported *cdk8*, *med12* and *med13* mutants which suggest that these four genes are likely to act together as part of the conserved CDK8 module. Besides, similarities in various phenotypes of *cycc1* mutants of *A. thaliana* and pea *late4* mutants (chapter 3) indicate that function of this novel flowering gene is likely to be conserved between these two plant species.

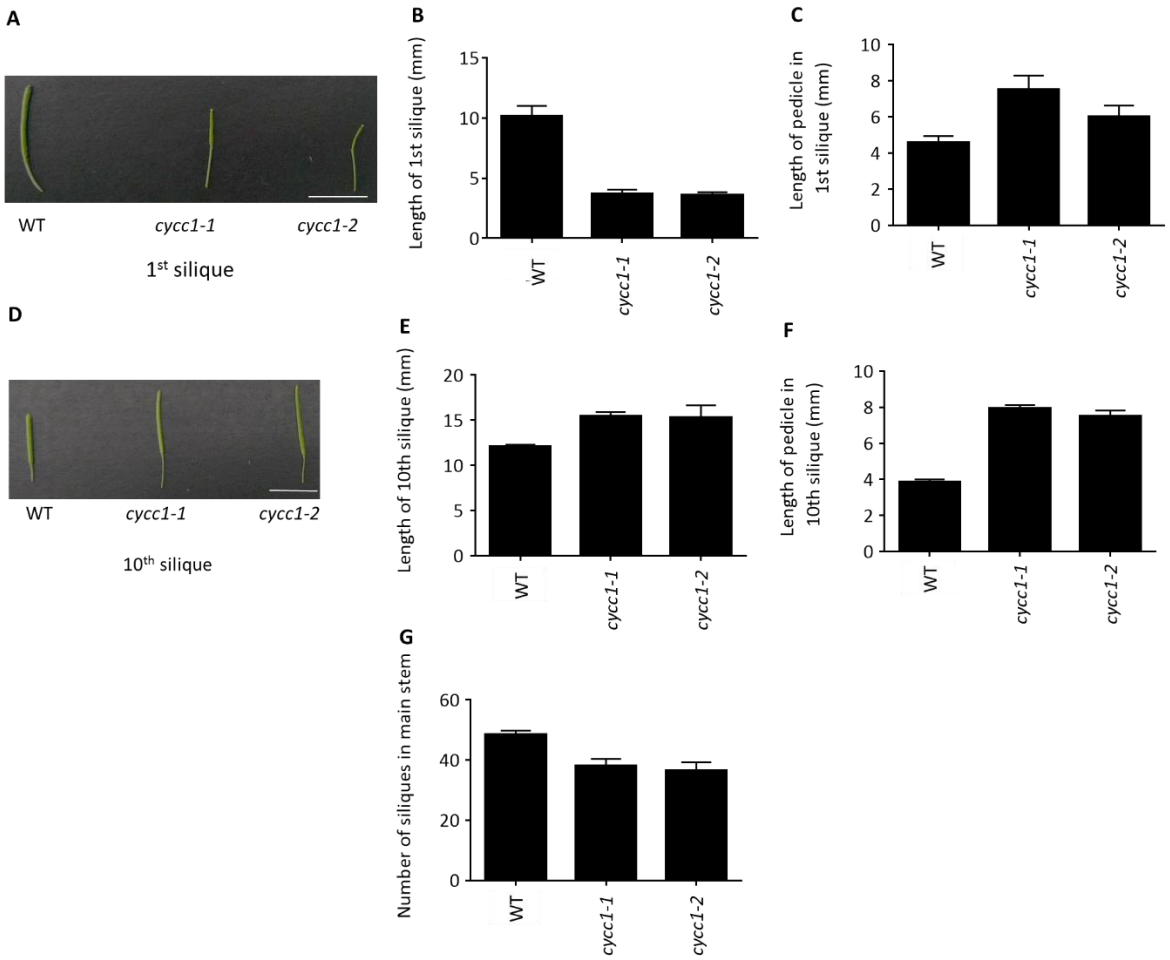


Figure 5.13. Alteration in different traits of siliques formed in the *A. thaliana* Columbia-0 (WT) and *cycc1-1* and *cycc1-2* mutants. **(A-C)** 1st silique, data represents mean ± SE for n = 5-8 plants. **(D-F)** 10th silique, data represents mean ± SE for n = 5-8 plants. **(G)** Number of siliques formed in the main stem (bolt) of WT and mutants, data represents mean ± SE for n = 5-8 plants. Scale bar represents 10 mm.

5.5 Discussion

5.5.1 LATE3 and LATE4 are pea homologues of mediator complex genes CDK8 and CYCC1 respectively

The multiprotein mediator complex was initially discovered in *Saccharomyces cerevisiae* due to its response to an activator of transcription (Flanagan et al. 1991; Kelleher et al. 1990) and it was later found to be involved in transcription of most of yeast genes (Holstege et al. 1998). This complex was then discovered in humans as well as other mammals (Berk 1999; Gu et al. 1999; Conaway et al. 2005). After that, it was discovered in the model plant species *A. thaliana* (Backstrom et al. 2007). Overall, the role of the mediator complex that has been described in all these systems is to regulate the transcription of wide range of genes by forming an even larger complex with general transcription factors (GTF) and gene specific transcription factors so that the RNA polymerase II can initiate the process of transcription. In human, it has been reported that the CDK8 module (CDK8, CYCC1, MED12 and MED13) of the mediator complex binds reversibly with the core mediator comprised of head, tail and middle module and thereby the process of reinitiation carried out by RNA polymerase II at the transcription initiation site is blocked which results in maintenance of optimum transcription of genes, and this process absolutely crucial for normal functioning of various biological processes (Knuesel et al. 2009a).

In the present study, pea mediator complex components *Cyclin dependent kinase* (*PsCDK8*) and *Cyclin C* (*PsCYCC1*) has been identified as candidates for *LATE3* and *LATE4* gene respectively via genetic mapping, RNA sequencing and phylogenetic analysis.

5.5.1.1 Potential effect of mutation in PsCDK8 of late3 mutants is rare alternative translation and most prevalent type of alternative transcription

In *late3-1*, site of EMS mutation was detected at -17 bp of the usual translation initiation site (ATG) at 5'UTR (Figure 5.6 A, Figure A2.1 A) leading to the introduction of a potential alternative start codon (GTG -> ATG). Two possible hypothesis could be considered to explain the effect of the *late3-1* mutation on the function of the protein. First hypothesis: the 40S small ribosomal subunit which engages with Met-tRNA and other translational initiation factors start the process of mRNA scanning at the 5' capped terminus until it encounters the first AUG codon, as a result the 80S ribosomal subunit can start the process of translation

(Kozak 1980, 1989, 1992, 1995). It has been shown that the sequences flanking the translation initiation site from -6 to -1 and +4 plays crucial role in determining the status of the scanning process (Kozak 1987, 1984a, c, b). This consensus sequence (GCCRCC**AUG**G) is known as Kozak sequence where R is represented by a purine base, i.e., G or A and variation may occur in all these bases. Presence of R in -3 position and G in +4 position is considered as strong consensus while adequate consensus should match atleast in one of these positions while a weak consensus do not have matching in any of the two positions. Overall strength of the translation is also found to be affected by the CC in -2 and -1 position as well as the G in -6 position (De Angioletti et al. 2004). Based on this understanding, it was found that the Kozak sequence of *PsCDK8* gene in the WT genotype, GCAACC**AUG**G fulfils all the aforementioned criteria at position -6, -3, -1 and +4 (Figure A2.1 A). In contrast, the new Kozak sequence, AACAAA**AUG**A introduced due to mutation at the 5'UTR of *PsCDK8* in *late3-1* mutant satisfies these rule only for -3 position. Due to this suboptimal Kozak sequence of the new start codon in *late3-1*, leaky scanning could follow where the ribosomal translational machinery would potentially bypass the first start codon and then initiate translation in the more suitable, second start codon. The hypothetical ORF starting at the alternative start codon would likely generate a protein of 25 aa long (Figure 5.6 E) which would be entirely different than the WT PsCDK8 protein as it would lack the major protein kinase domain that is needed for the function of the gene. So, if the entire translation would have occurred at this site, then a strong phenotype in terms of NFI was expected to occur similar to *late3-2* and *late3-3*. But, it was found that *late3-1* flowers at node 28 while the other two alleles flower at around node 42. This give hints that a certain proportion of the WT PsCDK8 protein is still produced in the *late3-1* mutant and it is likely that this may have occurred due to the leaky scanning where the translation machinery have probably evaded the first initiation site and preferentially started the process of translation at the second, stronger Kozak consensus sequence. Second hypothesis: it was reported previously that presence of upstream open reading frame at the 5'UTR (uORF) play role in the regulation of eukaryotic gene expression (Lovett and Rogers 1996; Vilela and McCarthy 2003; Somers et al. 2013; Wethmar et al. 2014). In a study, it was found that 50% of human genes possess 5' uORF which causes reduction in gene expression (Calvo et al. 2009). Depending on this idea, the new AUG at the 5'UTR of *PsCDK8* of *late3-1* may act as initiation site for a uORF. It has been reported that the translation of the actual start codon can only occur when the uORF translation ceases within a certain distance from

the original start codon so that the 40S ribosome can recruit the translation initiation factors and reinitiate the translation process. However, when the uORF overlaps the normal ORF as is the case occurred for *PsCDK8* in *late3-1*, then translational machinery is unable to get access to this internal ORF and its subsequent translation is suppressed (Kozak 2002). If the actual translation of *PsCDK8* gene would have occurred according to the uORF hypothesis, then it was likely that a very strong phenotype would have resulted due to total repression of the original ORF. Considering about these two possibilities and comparison of phenotypes in all the three *late3* mutant alleles and WT, the Kozak hypothesis seems to be the more likely scenario of downstream consequence of *late3-1* mutation (hypomorphic allele). This is assumed because the phenotype of *late3-1* was intermediate (partial phenotype) between WT and stronger *late3* alleles (i.e., *late3-2* and *late3-3*) providing hints that residual WT PsCDK8 protein activity is retained, rather than complete loss-of-function of the protein. A recent study involving similar mutation at 5'UTR (372 bp) in *SOX9* gene of human showed occurrence of campomelic dysplasia (CD) that is a semi-lethal developmental disorder (von Bohlen et al. 2017). In this study, the mutation G->A occurred at -185 bp which introduced a new start codon and translation at this site was hypothesized to produce a novel protein that terminates at +4 of normal start codon where the uORF would prevent access to the normal ORF. In contrast, the Kozak sequence around the new start codon was found to be stronger than the normal start codon. Generation of both novel and WT protein in the patient carrying mutation in *SOX9* gene was confirmed by in vitro transcription-translation and transient transfection studies. This result was consistent with the Kozak theory, therefore surviving CD patients were considered to carry hypomorphic rather than complete null allele of the *SOX9* gene.

For *late3-2* and *late3-3*, mutation occurred in *PsCDK8* gene at the 3' acceptor splice site of intron 12 (AG -> AA) and 5' donor splice site of intron 4 (GT -> AT) respectively (Figure 5.6 A, Figure A2.1 B-C). This caused skipping of exon 13 in *late3-2* and retention of 7 bp of intron 4 in *late3-3*. In a comprehensive study involving 43337 splice junction pairs, it was found that 98.7% of the annotated mammalian expressed sequence tag (EST) carries canonical dinucleotide GT and AG at 5' and 3' splice sites (Burset et al. 2000). These sites are important for recognition and initiation of the process of splicing of introns and according to the human genome mutation database, over 9% of the 141,000+ mutations that were published are

splicing mutation (Lewandowska 2013). The splicing process is regulated by a highly complex and multi-protein machinery called spliceosome (Nilsen 2003). Some of the key players in this regard are the five uridine rich small ribonucleoproteins (UsnRNPs) such as U1, U2, U4, U5 and U6 sRNA. Initiation of splicing occurs by complementary binding of the U1sRNP to the 5' splice consensus sequence (GTRAGT, R=G/A) which is accompanied by joint association of UA2F35 and U2AF65 proteins at the 3' splice site (CAG) and pyrimidine tract of the intron. Various proteins along with other UsnRNPs then complete the process of splicing of introns and joining the neighbouring exons (Black 1995; De Conti et al. 2013; Lewandowska 2013). Based on this understanding, it is highly likely that the underlying specific pea splicing machinery could not recognize the altered 3' splice site of intron 12, i.e., AA of *PsCDK8* gene in *late3-2*. In this case, the process of splicing that was initiated for intron 12 at the 5' splice site may have continued through exon 13 and it was probably terminated at the normal 3' splice site of intron 13 (Figure A2.1 B). It was probably due to this potential mechanism, the exon 13 was completely deleted as it was not recognized as an exon, rather part of long intron 12 in *late3-2.* Exon recognition also depends on exon splicing enhancer sequence (ESE) potentially located within exons that acts as the binding site of SR protein which then assists positively in the process of exon recognition by cross talk with other spliceosome machinery (Pertea et al. 2007; Long and Caceres 2009). In contrast, exon splicing silencers (ESS) which also could reside in the exon prevent exon recognition via binding with heterogenous nuclear ribonucleoproteins (hnRNPs) (Zhu et al. 2001; Bates et al. 2017). It might be a case for the exon skipping observed in *late3-2* where mutation in 3' acceptor site of intron 12 of *PsCDK8* have probably affected proper binding of the spliceosomes, and as a result any putative ESE residing in exon 13 of *PsCDK8* could not function properly or putative ESS may have been activated which ultimately resulted in skipping of exon 13.

As far as *late3-3* goes, the U1snRNP spliceosome may not have detected the modified 5' splice site, i.e., AT of intron 4 of *PsCDK8* leading to alternative splicing process where subsequent splicing occurred when this splicing machinery encountered the first possible 5' cryptic splice site located at +7 bp of intron 4 (Figure A2.1 C). This is the likely process through which exon 4 of *PsCDK8* has retained 7 additional bases in *late3-3* relative to WT. The cryptic splice sites (CSS) usually reside within exons and introns which also have the consensus splicing motif, but the relevant splicing machinery preferentially activates splicing at the authentic splice

sequence motifs because of the competitive advantage/strength over the CSS (Roca et al. 2003). The authentic , consensus 5' splicing motif of intron 4 of *PsCDK8,* **GTA**C**G**A has matching at four different positions with the consensus 5' splice motif GTAAGT whereas the 5' cryptic splice motif (CSM), **GT**TT**G**C activated in the intron 4 of *late3-3* genotype matches in three positions (Mount 1982). Hence, the weaker 5' CSM (also the first potential site) of intron 4 of *PsCDK8* was chosen in *late3-3* only under the circumstance of mutation occurring at the authentic consensus 5' splice site of the same intron.

Both the *late3-2* and *late3-3* seem to be null alleles as the hypothetical ORF of their CDK8 gene is likely to generate truncated protein eventually affecting the protein kinase domain needed for proper function of the gene (Figure 5.6 E). This is probably the reason because of which more severe phenotypes were observed in these mutants compared to *late3-1* where the PsCDK8 protein likely to have retained certain level of function.

5.5.1.2 Mutation in PsCYCC1 of late4 mutants generated rare types of splice variants

The present study revealed occurrence of EMS mutation for *late4-1* genotype at the 5th position of 5' splice motif of intron 5 in (GTAAGC->GTAAAC) *PsCYCC1* gene. Since this mutation was not in exon, so it required additional step of validating potential alternative splicing hypothesis via cloning and subsequent sequencing of different colony PCR products amplified around the site of mutation. As discussed in previous section, the consensus 5'splice motif (GTRAGT, R=G/A) is crucial for proper binding of U1snRNP which initiates the process of splicing. The +3 to +6 position of 5' splice motif (i.e., RAGT) are less conserved across various species than first two positions (i.e. GT, 5'SS) (Iwata and Gotoh 2011), however the 5th position, i.e., G is invariant in yeast (Lesser and Guthrie 1993). In a previous study involving 5,265 short introns (≤116 bp), G was found to be present in highest frequency compared to other three bases A, T and C (Lim and Burge 2001). Similar pattern of results was also found for 3,577 introns (≤60 bp) of *C. elegans*; 3,737 introns (≤81 bp) of *D. melanogaster*; 33,666 introns (≤134 bp) of human. In contrast, all of the 152 introns (≤191 bp) of *S. cerevisiae* possessed G in this 5th position. In the current study, most of the 50 analysed introns were found to carry G (58%) in the 5th position of 5' splice motif which provided hint that this position is likely to be highly conserved among eukaryotic introns. Most importantly, all the eight introns of *PsCYCC1* including the short intron 5 (139 bp) possess G in the 5th position of 5' splice motif which gives hint that a particular type of spliceosome machinery might be

initiating the process of splicing of WT *PsCYCC1* pre-mRNA in pea by specifically recognizing this 5th position. It can therefore be assumed that mutation at this non-canonical position for *PsCYCC1* gene have affected the normal splicing process which led to the generation of five different splice variants as the splicing machinery scanned across the pre-mRNA to look for 5' cryptic splice motifs (CSM) around the site of mutation. The WT 5' splice motif of intron 5, i.e., GTAAGC of *PsCYCC1* matches the consensus motif, i.e., GTRAGT in the first five bases and this is probably the reason because of which the pea splicing machinery preferentially picks this motif to initiate splicing of intron 5 even though there are some potential cryptic splice sites embedded in nearby exons and introns.

For *late4-2*, presence of non-sense mutation or premature stop codon (CAG->TAG) in exon 5 of *PsCYCC1* was found which also resulted in generation of four different types of alternative transcripts (Figure 5.9 F, Figure A2.2 A-D). Eukaryotic organisms have evolved a conserved surveillance mechanism through which it scans for presence of such premature stop codons (PTC) which then undergoes a process called non-sense mediated mRNA decay (NMD) in order to get rid of such aberrant transcript which is likely to generate a truncated protein having deleterious effect on the organism (Maquat 2005; Brogna and Wen 2009; Lykke-Andersen and Jensen 2015). NMD usually takes place when the PTC is present more than 50 bp from the 3' end of the respective exon. Since the PTC in *late4-2* was observed -37 bp from the 3' end of exon 5 of *PsCYCC1*, so it is unlikely that this mRNA would be subjected to NMD. In the current study, vast majority of the transcripts, i.e., transcript type 1 (81%) of *late4-2* was similar to WT apart from the position where the mutation occurred. Any potential protein would be truncated with impact on cyclin N domain and these proteins may have had strong harmful effect on the plant resulting in strong malformed phenotype. On the other hand, rest of the 19% transcripts, i.e., transcript type 2-4 were generated via various patterns of alternative splicing. It has been proposed in a previous study that the PTC could influence selection of 5' SS through surveillance so that modulated transcripts would not include the stop codon (Li et al. 2002). Analysis of the sequence of the exon 5 revealed presence of putative exon splicing silencer motifs (ESS) introduced in *late4-2* due to mutation which may have caused the skipping of this PTC carrying exon in transcript type 2 (Figure A2.2 B). Exon skipping associated with non-sense mutation have been reported previously in different studies (Bach et al. 1993; Dietz et al. 1993; Gibson et al. 1993). Transcript type 3 which showed

partial intron 5 retention underwent splicing at 5' CSM (GTGAGT) located within intron 5 which exactly matched consensus 5' splice motif (Figure A2.2 C). It has been reported that presence of in-frame stop codon prior to such 5' CSM hinders the splicing machinery to pick this over the authentic 5' splice motif which might be less strong (Li et al. 2002). Interestingly, a potential stop codon (TAA) was found to be present in-frame at +17 bp of intron 4 because of which the subsequent stronger 5' CSM was probably not selected initially. But, the non-sense mutation in exon 4 may have affected this suppression process resulting in activation of the stronger 5' CSM in intron 5. Intron retention due to non-sense mutation was also reported previously in different studies (Lozano et al. 1994; Maquat 1995). In case of transcript type 4, the activation of a 5' CSM in exon 4 just before the PTC would probably mean that the underlying surveillance machinery acted in creation of an exon without the presence of this PTC (Figure A2.2 D). All the proteins that are likely to be generated in *late4-2* (null allele) would be non-functional and lack of production of any WT proteins is probably the reason why this mutant allele exhibited stronger phenotypes than *late4-1*. Thus, the present study shed light on the complex mechanism that is associated with generation of de novo splice variants as an additive effect of non-sense mutation.

In case of *late4-1*, certain proportion of WT protein of *PsCYCC1* is likely to be produced by the 18% of the transcript belonging to type 1 which had similar transcript sequence like WT (Figure 5.9 G). This would mean that the splicing machinery probably chose the mutated 5' splice motif, GTAAAC (WT, GTAAGC) to splice out intron 5. For transcript type 2, splicing occurred at a 5' CSM (GCTAGG) located in exon 5 (Figure 5.9 G, Figure A2.2 E) which matches the expected 5' consensus splice motif in three out of six positions, but importantly it picked the site having G in 5th position suggesting the importance of this position for recognition by relevant splicing machinery. For transcript type 3 (Figure 5.9 G, Figure A2.2 F), the entire exon 5 was found to be missing which may have occurred because of the failure of splicing machinery to initiate splicing of intron 5 at the mutated 5' splice motif, and as a result splicing process that started at the 5' authentic splice motif (having G in 5th position) of intron 4 continued through exon 5 and splicing ended at the 3' splice site of intron 5. Alternatively, putative exon splicing silencer motifs (ESE) of exon 5 was probably unable to interact with respective SR proteins and surrounding spliceosomes due to the mutation and therefore exon 5 was not recognized being a bona fide exon. Next, transcript type 4 was found to carry a

cryptic exon named as exon 5a which was embedded within intron 4 and interestingly this cryptic exon was bordered by a 3' cryptic splice site (CSS) and a 5' CSM (GTAAGA, stronger than the mutated 5' consensus splice motif) having G in 5th position (Figure 5.9 G, Figure A2.2 G). Occurrence of such cryptic exon was found previously specially in the case when a 5' CSM matches the consensus 5' splice motif more than a mutated 5' splice motif (Rogan et al. 1998; Wessagowit et al. 2005). So, the splicing machinery probably recognized the 3' SS as well as stronger 5' CSM of exon 5a due to mutation at normal 5' SS of intron 5 (resulting in weaker consensus splice motif) and splicing occurred at either side of this cryptic exon. Finally, transcript type 5 was found to retain the entire intron 5 producing a large exon 5 which may have occurred due to non-recognition of the mutated 5' splice motif of this intron by spliceosome (Figure 5.9 G, Figure A2.2 H). This type 5 splice variant also showed presence of cryptic exon 5a as explained earlier for transcript type 3. It is likely that other types of splice variants may have also been generated which were not detected via the annealing temperature of the PCR or was not included within the 31 colonies that were used for this study.

The production of certain proportion of WT protein is probably the reason why *late4-1* had a slightly milder phenotype than *late4-2* which flowered at node 38 and 40 respectively. In a previous study in human, mutation at the 5th position of intron 32 in dystrophin gene was also found to be associated with generation of rare splice variants ultimately causing a serious type of muscular human disease called Duchenne muscular dystrophy (Thi Tran et al. 2005). Overall, the aforementioned results from the present study unveiled the potential reliance of splicing machinery on the 5th position, i.e., G of 5' splice motif in a higher species like pea and how dramatically mutation at this site could affect the protein function.

5.5.2 Strong conservation of mediator complex components in M. truncatula and pea

In order to expand the horizon further beyond *CDK8* and *CYCC1*, the components of the broader mediator complex in pea and *M. truncatula* were identified in this study (Table 5.4). Studies regarding mediator complex in higher plant species is still at a preliminary stage. It has been proposed previously that the head, tail and middle module forming core mediator remained conserved in animal/plants since its origin as a 17 subunit proto-complex 1-2 billion years ago in protest (Bourbon 2008). The CDK8 kinase module consisting of CDK8, CYCC1, MED12 and MED13 is the result of latest evolutionary event. In this direction, independent

lineages were evolved for lower eukaryotes like *S. cerevisiae* and mammals/plants where potential duplication of different subunits have occurred. So far, the mediator apparatus has been identified in the model species *A. thaliana* and few other dicots and monocots (Backstrom et al. 2007; Mathur et al. 2011). These studies revealed conservation, deletion and duplication of different fungal/animal mediator complex components in plants along with identification of plant specific components such as MED34, MED35, MED36 and MED37. Based on this information, the findings of the present study revealed conservation of most of the subunits between the three species. However, supplementary copies of some of the subunits were found to be in one species only or two out of the three species.

Taken together the results of current study involving two legumes and those involving various monocots and dicots carried out by Mathur et al. 2011, it could be stated that the mediator complex is deeply conserved among the members of angiosperms. The divergence that has inevitably took place for this very important plant protein complex among these species was probably to facilitate individual functioning and survival. Thus, the present study has laid a foundation for future investigations which could unveil the function of each of these subunits which would enable the scientific community to strengthen knowledge and understanding about the distinct process of transcriptional regulation and in a broader sense evolutionary adaptation mechanism in pea as well as other legumes.

5.5.3 Function of CYCC1 gene is likely to be conserved in A. thaliana and pea

The key feature of the two insertion mutants of *AtCYCC1* namely *cycc1-1* and *cycc1-2* as observed in this study was their significantly delayed flowering relative to WT Col-0 under LD condition (Figure 5.12 A-B). In previous studies involving two other CDK8 module components *AtMED12* and *AtMED13*, the respective mutants namely *crp-2* and *mab2-1* generated in the Landsberg eracta (Ler) background also showed delayed flowering under LD conditions compared to WT (Imura et al. 2012). In another study involving *cdk8*, *med12* and *med13* mutants of *A. thaliana,* late flowering phenotypes were observed for all these three mutants where the mutant phenotype of *med12* and *med13* were stronger than *cdk8* (Zhu et al. 2014). Similar studies were also performed on other components of the mediator complex. In such a study involving *pft1* mutant of *AtMED25* also demonstrated very late flowering phenotype under LD (Cerdán and Chory 2003; Kidd et al. 2009). Likewise, *med18* and *med8 mutants* in *A.*

thaliana flowered considerably late compared to WT in LD in two independent studies (Kidd et al. 2009; Zheng et al. 2013).

Apart from flowering time, alteration in other developmental traits such as increased branching, stem height as well as number and structure of silique (Figure 5.12 C-D, 5.13 A-G) were also observed in the aforementioned T-DNA lines of *AtCYCC1*. Among these, the thinner and longer silique that was found in these mutants compared to WT corresponds to similar structure of the insertion mutants *crp3* and *crp4* of *AtMED12* (Imura et al. 2012). Even though details of flower morphology were not studied in the current project, however modification in shape of flower was observed and a proportion of the seeds derived from thinner siliques of the *cycc1-1* and *cycc1-2* mutants are likely to have impact on germination because reduced germination rate was observed for the ordered seeds. Mutant for *AtCDK8* namely *hen3* has been studied before and its role in floral and vegetative organ identity such as leaf was proposed (Wang and Chen 2004). Besides, positive role for all the four components of *A. thaliana* CDK8 module against fungal disease caused by *A. brassicicola* was reported previously where *cdk8, cycc1, med12* and *med13* mutants developed similar phenotypes, i.e., increased lesion in leaf upon infection by the aforementioned fungus (Zhu et al. 2014). Moreover, global regulatory role of *CDK8*, *MED12* and *MED13* in *A. thaliana* have been proposed through different studies (Ng et al. 2013; Gillmor et al. 2014). With regard to this, *AtCYCC1* may have similar type of universal regulatory role and this was revealed to a certain extent by similarities in flowering phenotype of *cycc1* mutants with that of already known mutants (*cdk8, med12* and *med13*) of the CDK8 module. Moreover, similarities in the results of this chapter and those of pea *late4* mutants presented in chapter 3 also provide compelling evidence about the conservation of the underlying function of the *CYCC1* gene across lower and higher plant species. Furthermore, it can be speculated that various components of the CDK8 module in plants such as pea are likely to act through mutual interaction as known in other systems such as yeast, *Drosophila* and human (Loncle et al. 2007; Tsai et al. 2013; Wang et al. 2013). More discussion in this direction will be carried out in the following chapters.

5.5.4 RNA sequencing proved to useful for candidate gene analysis in a species like pea, but coverage is important

In order to select candidate gene(s) from genetic mapping studies in plants, the RNA sequencing technology is making rapid inroads. For instance, it has been exploited for

detecting EMS induced causal mutations in *glossy 3* (*gl3*) and *glossy13* (*gl13*) genes of maize (Li et al. 2013b; Liu et al. 2012). Likewise, fine mapping and global transcriptome analysis unravelled candidate genes involved in hybrid lethality in cabbage (Xiao et al. 2017). In addition, candidate genes likely to be involved in salt tolerance at the seedling stage in wild rice was selected upon combined RNA sequencing and QTL mapping (Wang et al. 2017). Moreover, mapping and subsequent selection of candidate gene responsible for female-male sterility in rapeseed was done via RNA sequencing (Teng et al. 2017).

As RNA sequencing is only useful for identifying polymorphism within the transcript sequence, therefore this method was actually of not much use for the intron based 5' splice site motif mutation that was found in *late4-1* by direct sequencing. Besides, it would not have also been useful had *late3-2* or *late3-3* been chosen as the representative allele for cDNA library preparation and subsequent RNA sequencing run for *late3* mutation since both these genotypes were also found to have splice site mutations. Nevertheless, higher coverage second run involving *late3-1* markedly increased the number of reads as well as coverage for almost all the 62 transcripts. Based on this data, putative functional polymorphism was detected in the 5'UTR of pea *Cyclin dependent kinase 8* (*CDK8*) gene in *late3-1* genotype within the defined interval of *PsLGIII* (Figure 5.4 A-B).

Depending on the experience of this study few suggestions could be given for future usage of this technology for similar purposes. Firstly, selection of wide range of tissues from different developmental stages is likely to increase the detection of a transcript that is expressed at low level specially keeping in mind that the underlying gene of an EMS mutant could sometimes be distantly related to flowering and published information on such genes may be limited. Secondly, coverage of a particular run is an important matter. In addition, potential likelihood of rare mutation events such as those occurring in intron/promoter regions need to be kept in mind and it could be considered whether WGS can be carried out. In this regard, genomic DNA (gDNA) of the specific chromosome (defined by genetic mapping) could be sequenced if flow sorted chromosome separation (Doležel et al. 2012) is possible for the species of interest. Likewise, targeted NGS could be performed to sequence the gDNA within the chromosomal region of interest (Summerer 2009; Mertes et al. 2011). Moreover, to make better use of the cost related to RNA sequencing while taking advantage of this technology, consideration could be given to design bulk segregant RNA sequencing mapping experiments

at the beginning (Du et al. 2017; Liu et al. 2012). This is a powerful method through which genetic mapping (done by quantification of allelic frequencies of genetic markers between two different sample pools that differ for the phenotype of interest), potential identification of causal mutation within differentially expressed candidate genes and downstream analysis about the effect of the candidate gene on expression of global transcriptomics could be done by using the data generated from the same run with no additional cost. In such case, even if the potential causal mutation is not found in the transcript sequences, the generated data gained through a certain cost would still be very useful for the other two aforementioned purposes.

5.5.5 Phylogenetic analysis suggests that the function of members of SET domain proteins is likely to be conserved in pea

Various members of the SET domain proteins/methyltransferases in *A. thaliana* were reported previously to be involved in the process of flowering, determination of cell fate, floral organogenesis, leaf morphogenesis and seed development (Pien and Grossniklaus 2007; Shen and Xu 2009; Berr et al. 2011). The SET domain protein carries out histone modification via mono, di, tri methylation is carried out by methyltransferases (Tschiersch et al. 1994). Based on this understanding and subsequent identification of two pea transcripts PsCam038017 and PsCam029108 (located within/near the defined interval of *LATE3* and *LATE4* genetic map respectively) as the *A. thaliana* orthologues of *AtSDG15* and *AtSDG10* through phylogenetic analysis led to further studies of the aforementioned pea genes as candidates. Since the sequence analysis did not reveal any functional polymorphism (Figure 5.2 A-B), so both the genes were ruled out as the likely candidates. Nevertheless, the phylogenetic tree developed in the process serves as a pioneer work on the SET domain proteins in pea and could assist in getting insight into the role of these proteins in different biological process in pea and *M. truncatula*.

In the current study, members of five different SDG protein families in *A. thaliana* carrying full SET domain as proposed by Springer et al., 2003 were included for the identification of pea and *M. truncatula* orthologues (Table 5.3). Overall, *A. thaliana* SDG proteins were found to be conserved in both the legumes in the current study. Since the plant SDG family carries more members than animals (e.g. 14 and 17 in *Drosophila* and mouse respectively) as also revealed in the current study, therefore it was proposed previously that SDGs underwent

various round of duplication prior to divergence from animals (Springer et al. 2003). It was probably due to polyploidization/chromosome addition, reorganizations, small scale transposition, illegitimate recombination or retransposition (Springer et al. 2003). Two other feature which was reported is the substrate specificity of various groups of SDG proteins and likelihood of them working in a complex. Members of the group I SDGs were reported to be involved in H3K27me2/3, group II in H3K36me2/3, group III in H3K4me2/3, group IV in H3K27me1 and group V in H3K9me2 (Jackson et al. 2002; Alvarez-Venegas et al. 2003; Lindroth et al. 2004; Xu et al. 2008; Jacob et al. 2009). In *A. thaliana*, methylation of H3K9 and H3K27 has been associated with transcriptional repression whereas H3K4 and H3K36 promote gene activation in plants and animals (Martin and Zhang 2005; Liu et al. 2010). Thus, SDGs from group I, IV and V act in repression of gene function whereas group II and III are involved in positively regulating gene function. The higher level of conservation in the structure of functional domains (shown for group I and IV in appendix sequence alignment A2.1, A2.2) of orthologous proteins in *A. thaliana*, *M. truncatula* and pea suggest that the function of gene activation and repression by various methyltransferases is probably conserved among the three species. Future studies involving individual methyltransferases from these two legumes could unveil function of these important plant development regulators.

5.5.6 Concluding remarks

In the present chapter, *LATE3* and *LATE4* were identified as pea homologues of *CDK8* and *CYCC1* respectively which are component of the highly conserved eukaryotic global transcription regulatory component called mediator complex. Thus, this work further strengthened the hypothesis that was made in chapter 3, i.e., *LATE3* and *LATE4* are universal regulatory genes and they are likely components of the same regulatory pathway in pea. It laid proper ground for further validation in this direction about the function of the *PsCDK8* and *PsCYCC1* through genetic and regulatory interaction with known pea flowering genes. In addition, studies on very rare and novel form of splicing events in *late3* and *late4* mutants opened up a new window of information on the complex mechanism of alternative splicing in a higher plant species like pea. Moreover, detailed phenotypic characterization of the mutants of *CYCC1* gene in *A. thaliana* similar to previously reported three members of CDK8 module, i.e., *CDK8*, *MED12* and *MED13* as well as those of *LATE4* gene in pea of the present

study indicated conserved function of CDK8 module subunits in pea. Furthermore, identification of various components of the SET domain proteins as well as entire mediator complex in pea and *M. truncatula* revealed overall conservation of both the components and these could be used in future as an important resource to understand the underlying molecular mechanism driving various developmental processes in pea as well as other legumes.

Chapter 6: Analysis of various interactions for *LATE3* and *LATE4* loci

6.2 Introduction

6.1.1 Relevance of genetic, regulatory and physical interaction experimentations

Genetic interaction between two non-allelic or different genes is often used in plant science in order to get an insight into the type of functional relationship that is prevailing between those two genes. Investigation on the genetic interaction between the two genes is carried out usually by studying the phenotype of the double mutants generated from single mutants of the respective genes. In the case when the double mutant exhibits phenotype similar to one of the single mutant rather than the other one, then it can be explained that the former gene is epistatic to the later one. The later gene in this case is termed as hypostatic to the former one. The epistatic gene is considered to act downstream of the hypostatic gene in a classical regulatory pathway/ switch regulatory pathway whereas the order is opposite in case of a classical assembly pathway/substrate dependent pathway (Michels 2002; Huang and Sternberg 2006). Such kind of understanding is imporant to organize a set of unknown genes in a regulatory pathway and the order in which they might act upon each other (Jarvik and Botstein 1973; Botstein and Maurer 1982; Avery and Wasserman 1992; Huang and Sternberg 1995). This in turn can be an useful tool to reveal the dynamic regulatory events underlying the biological phenomenon of interest.

The term genetic interaction is different than physical interaction where two proteins or protein/DNA (e.g. interaction between TF and target gene) interact directly. This is usually performed by checking interaction of two proteins where binding of one protein to the other would mean that they have physical interaction (Jones and Thornton 1996; Westermarck et al. 2013). Genetic interaction between two genes could mean direct impact on the function of the downstream gene by the upstream gene through co-operation with transcription factors that are involved in regulating the function of the former gene. This could be revealed by appropriate functional validation of observed physical interaction studies of proteins from the two genes. In contrast, genetic interaction could also be indirect resulting in modulation

in the expression patten of downstream genes. In this regard, subsequent research steps involving gene expression studies could enhance perception of the nature of underlying molecular consequences of the genetic interaction between the two genes. Such studies form the basis of addressing questions regarding regulatory role of a particular gene.

Based on these principles, genetic, regulatory and physcial interaction studies have played very curcial roles in unravelling the order as well as the nature in which various genes work in the flowering pathway in pea. Such studies have framed the basis of functional characterization of different early and late flowering mutants in pea which ultimately strengtened understanding how a given gene functions in controlling flowering in pea (Hecht et al. 2007; Liew et al. 2009; Hecht et al. 2011; Liew et al. 2014; Sussmilch et al. 2015; Ridge et al. 2016). Similar attempts have been made for characterization of *late3* and *late4* mutants prior to the initiation of the current study when the molecular nature of *LATE3* and *LATE4* was not known. As discussed in chapter 3, the similarities in phenotype between *late3* and *late4* mutants led to initial apprehension that these two genes probably act in the same pathway in a mutualistic dependent manner. In order to address this issue, *late4-2* and *late3-1* were crossed where non-additive effect of both these mutations in suspected double mutants was observed and the F_2 generation showed an approximate segregation ratio of 9 wild type : 7 mutant phenotype ratio (Weller et al., unpublished data). This observed alteration in the expected Mendelian F_2 segregation of 9:3:3:1 could have derived from duplicate recessive mutations where mutation in one gene affect the function of the other and such kind of genes involved in an epistasis relationship are termed as complementary genes (Miko 2008). Identification of *LATE3* and *LATE4* as *PsCDK8* and *PsCYCC1* (chapter 5) which are components of the highly conserved CDK8 module of the eukaryotic mediator complex strengthened perception about likely mutual interaction of these genes to carry out the same regulatory function. Therefore, it is worthwhile to confirm these results and conduct futher genetic interaction studies between *late3* and *late4* mutants as well as check physical interaction between the proteins encoded by these genes.

Two early flowering mutants *lf* and *det* were also used in preliminary investigations before the commencement of the present study in order to reveal their genetic interaction with *late3* and *late4* mutants (Weller et al, unpublished data). The *lf* mutant was reported initially as early flowering and sensitive to flowering singals which initiate flowering in a day neutral

manner, but show photoperiod dependency for other traits (Murfet 1971, 1975; Murfet 1993). On the other hand, *det* mutation resulted in termination of shoot meristem activity once flowering is initiated (Swiecicki 1987; Singer et al. 1990; Murfet 1993). Later on, *LF* and *DET* were identified as the pea homologs namely *TFL1c* and *TFL1a* of *A. thaliana TFL1* gene which in the model plant system controls both determinacy as well as flowering (Foucher et al. 2003). The *tfl1* mutant in *A. thaliana* is early flowering and determinate. Due to evolutionary divergence, this dual function of *TFL1* gene have been allocated to two different genes in pea where *LF*/*TFL1c* regulates initiation of flowering by blocking vegetative to inflorescence meristem termination. In contrast, *DET*/*TFL1a* mediates determinacy by repressing inflorescence meristem to floral meristem transition. As both of these mutants along with *late3* and *late4* exhibit distinct flowering phenotype, therefore genetic interaction between them are likely to reveal better understanding about the mechanism of functioning of *LATE3* and *LATE4* genes. Initially*, late3-2* and *late4-1* were crossed with *lf* to generate *late3-2 lf* and *late4-1 lf* double mutants (Weller et al, unpublished data). In a separate cross, *lf* was crossed with *det* which led to identification of *lf det* double mutants. After that, *late3-2 lf* and *late4-1 lf* double mutants were crossed with *lf det* double mutants and subsequent identification of putative *lf det late3-2* and *lf det late4-1* triple mutants based on phenotype was also made from the segregating population. Further confirmation of the phenotypes of potential double and triple mutants led to usage of these available resources during the present study.

Similarly, another early flowering mutant *sn* was used for understanding the underlying genetic interation with *late3* and *late4* before the launch of the current study (Weller et al., unpublished data). Prior to the molecular era, *sn* mutant was reported to flower in a day neutral fashion and have positive role in transition from two to more than two leaflets per leaf (Barber 1959; Murfet 1971; Murfet 1993). Functional characterization of the *sn* mutant later led to the identification of the underlying gene as the pea homologue of *A. thaliana* circadian clock gene *LUX ARRHYTHMO* which is a GARP transcription factor and role of this gene in regulating expression of pea *FT* genes was established (Liew et al. 2014). In order to check genetic interaction, *late3-2, late4-1* and *late4-2* were crossed with *sn-4* and putative double mutants were found within the segregating population based on phenotype (Weller et al., unpublished data). For the purpose of confirming the previously observed phenotypes of these double mutants, these resources were also used in the current study.

Among the five *FT* genes in pea, *FTb2* which is expressed only in leaves was found to possess the features of classical mobile signal element "florigen" and its induction is associated closely with the physiological stimulus needed for flowering (Weller and Ortega 2015). Another *FT* gene *FTa1* was also reported as a likely contributor in the direction of generating mobile floral signal (Hecht et al. 2011). However, upregulation of *FTb2* was found to precede that of *FTa1* as needed for flower initiation in pea. A third important *FT* gene namely *FTc* is expressed only in the apex and it probably act in integration of floral signal from the leaf bound *FT* genes (Hecht et al. 2011). Under LD, *LATE1*/*PsGI* promote *FTb2* expression whereas *DNE/PsELF4* negatively regulate expression of both *FTa1* and *FTb2* under SD in leaf (Hecht et al. 2011; Laurie et al. 2011). Besides, *FTa1* is considered to be important for vernalization response in pea (Weller and Ortega 2015). The process of inflorescence development in pea is associated with three sequential events starting with conversion of vegetative meristem (VM) to indeterminate primary inflorescence meristem (I_1M) which in the next stage gives rise to determinate secondary inflorescence meristem (I_2M) and in the final stage floral meristem (FM) is derived leading to flower formation (Sussmilch et al. 2015). Various genes act as marker for each of these stages such as pea VM identity gene *LF* (*PsTFL1c*) (Foucher et al. 2003), I_1M specific gene *DET* (*PsTFL1a*) (Foucher et al. 2003), I_2M specific *VEG1* (*PsFULc*) (Taylor et al. 2002), FM identity gene *PIM* (*PsAP1*) (Taylor et al. 2002) and *UNI* (*PsLFY*) (Hofer et al. 1997) and overall inflorescence fate determinant specific gene *VEG2* (*PsFD*) (Sussmilch et al. 2015). Therefore, regulatory interaction studies of these important pea flowering genes in the flowering mutants and wild-type have also been a key component in previous studies to understand how the known genes are affected by the mutation (Hecht et al. 2007; Hecht et al. 2011; Sussmilch 2014; Ridge et al. 2016). It also assists in getting insight into the type and level of role the identified gene (s) play in relation to regulation of known genes which ultimately is involved in proper maintenance of flower pathway in pea. However, research in this direction was not carried out prior to the commencement of the present study for *late3* and *late4* mutants.

6.1.2 Importance and applicability of systems biology based interaction network studies

Usage of genomics, proteomics, transcriptomics and metabolomics information have long been used by scientists in order to understand various biological processes. Systems biology deals with dynamic interaction between various components of the biological system using

network/models where each one of them function individually and thereby regulate the phenotype (Yuan et al. 2008a; Sheth and Thaker 2014). An interaction network consists of "nodes" which are the components of the system and "edges" where interaction between two nodes is illustrated by connecting lines (Sheth and Thaker 2014). Nodes are represented by proteins, genes, cis-elements, metabolites whereas edges depict the type of biological relationship between nodes such as physical, genetic and regulatory interactions, metabolic relationship, co-expression, co-localization, protein homology and algorithm based prediction. Such computational research helps to derive hypothesis which can then be validated by experimental research in order to generate in-depth knowledge about different signalling pathways in living system (Kumar 2013).

There are four main types of interaction networks are used in systems biology. Firstly, a gene-to-metabolite network exhibit regulatory interaction (edges) between genes and metabolites (nodes) (Sheth and Thaker 2014). Secondly, a protein-protein interaction network represents the direction of interaction (edges) between proteins (nodes) if known or non-direction if there is no evidence about the directivity of interaction (Sheth and Thaker 2014). Thirdly, regulatory interaction (edges) between transcription factors and their target genes (nodes) are illustrated in a transcriptional regulatory network where interaction could be at transcriptional or translational level (Yuan et al. 2008a; Sheth and Thaker 2014). Lastly, nodes in a gene regulatory network are represented by genes, messengers or proteins whereas edges show different types of interaction such as activation, inhibition, repression or other type of functional interaction between the nodes (Sheth and Thaker 2014; Yuan et al. 2008a).

In human biology, co-functional interaction networks have been used previously for prioritizing causal genes using genome wide association studies (Taşan et al. 2014). Besides, gene regulatory nework are exploited as a blueprint or map of molecular interactions based on which novel biological hypothesis can be derived and it can be further investigated by using wet lab experiments such as gene expression and Chip-chip (Bussemaker et al. 2001; Basso et al. 2005). Since the challenges for plant scientists to meet the demand of global food, fodder and energy is increasing in an altered climatic conditions, so various interaction network is regulary being used by them in the modern era to address issues like bettering yield, pest management, abiotic stress tolerance, biofuel production and other relevant issues (Dortay et al. 2008; Morsy et al. 2008; Tai 2008; Yuan et al. 2008a; Yuan et al. 2008b; Ding et al. 2009).

A key protein-protein association network that is widely used is STRING (Szklarczyk et al. 2015; Szklarczyk et al. 2016). The aim of this database is to provide a system-wide knowledge about cellular functions by compiling information from various sources (e.g., published experimental results) on direct physical interaction and indirect functional interaction between expressed proteins which are biologically meaningful. As an example, the STRING database has been used previously for identifying functional interaction network partner of MED12 in human (Banaganapalli et al. 2016). Another such database is the genomde wide functional network called AraNET which was developed by incorporating various functional genomics, proteomics and comparative genomics databases in order to deliver inoformaiton on gene-to-phenotype association (Lee et al. 2010; Lee et al. 2015; Lee and Lee 2017). This database can aid in finding most likely new members of a pathway or phenotype of interest. Besides, it can also infer function of a query gene in terms of gene ontology biological process (GO-BP) annotations by analysing most enriched GO-BP terms obtained from network neighbours of the query. For instance, this database has been succesfully used in identifying new member of the leaf initiation pathway in non-model species such as maize (Lee et al. 2015). A third database namely PlantTFDB provides extensive, high quality resource for getting knowledge about regulatory interactions of plant transcription factors (TFs) and their target genes (Guo et al. 2008a; Jin et al. 2014; Jin et al. 2017). It can provide information on the predicted TFs regulating the function of a query gene by analysis the DNA binding motifs of large collection of TFs.

As revealed and discussed in the previous chapter, *LATE3* and *LATE4* was identified as pea orthologoues of *A. thaliana CDK8* and *CYCC1* which are component of CDK8 module belonging to the larger eukaryotic mediator complex that regulate transcription of various genes involved in different biological pathways (Yang et al. 2016). Due to the importance of CDK8 module in relevance to biological system mangement and lack of information about it in plant system, it is very reasonable to use STRING and thereby get insight about the broader network of genes that are likely to act via biologically relevant interactions in regulating the function of the *CDK8* and *CYCC1* genes in the model species *A. thaliana*. Such knowledge is likely to assist in developing hypothesis in *M. truncatula* and pea also. As the AraNET includes *M. truncatula*, therefore getting information on the gene-to-phenotype association for *MtCDK8* and *MtCYCC1* is likely to assist in cross checking the observed phenotypes of *late3* and *late4*

mutants in pea as well as those of *cycc1* mutants in *A. thaliana.* In addition, usage of the available resource of PlantTFDB database could generate crucial information on the TFs that are most likely to regulate the function of *MtCDK8* and *MtCYCC1* genes and this information can then be exploited to identify potential orthologues in pea.

6.2 Chapter aim

The main goal of this chapter was to develop knowledge on how *LATE3* and *LATE4* interact to regulate flowering in pea. To this end, validation of hypothesis from previous chapters that *LATE3* and *LATE4* act in the same regulatory pathway was carried out by phenotypic and genotypic analysis of *late3-1 late4-2* double mutant population. Besides, yeast two hybird studies was conducted in order to check whether CDK8 and CYCC1 proteins interact physically or not. In addition, role of *LATE3* and *LATE4* in mediating flowering was examined through further molecular physiological studies such as genetic interaction and expression analysis. As part of that, regulatory interaction of *LATE3* and *LATE4* genes with already known and important pea flowering genes was checked through Q-RT-PCR. Likewise, genetic interaction studies of late flowering *late3* and *late4* mutants with early flowering mutants such as *lf*, *det* and *sn* was performed via phenotypic characterization of relevant double and triple mutants. Moreover, in order to determine whether *LATE3* and *LATE4* act in the photoperiod and vernalization mediated flowering pathway, response of *late3* and *late4* mutants towards these factors were investigated. Furthermore, likely pattern of interaction between the various components of CDK8 module in *A. thaliana* have been revealed through generation of interaction network by using the web platforms namely STRINGv10.5. Similarly, components of a broader network of AtCDK8, AtCYCC1-1 and AtCYCC1-2 was developed using the same platform. Additionally, predicted functions of *CDK8* and *CYCC1* genes in *M. truncatula* was generated using AraNETv2.0 database in order to compare already obtained knowledge from the closely related species, i.e., pea. Last but not least, projections regarding the type of TFs regulating function of *CDK8* and *CYCC1* gene in *M. truncatula* have been made by using PlantTFDB4.0 database and orthologous genes for these TFs were identified in pea as well.

6.3 Materials and methods

Various materials and methods used specifically for this chapter are mentioned here. The general materials and methods exploited in this chapter are given in chapter 2. Details of sequences are provided in appendix 3.

6.3.1 Experimental set-up for genetic interaction studies

Plants for various genetic interaction studies were grown at the glass house of School of Natural Sciences, University of Tasmania at 20° C (18 hour photoperiod). The growth condition of the glass house has been mentioned in section 2.2. As mentioned earlier, Weller et al. (unpublished data) initiated genetic interaction studies between *late3* and *late4* mutants and observed a 9 wild-type : 7 mutant segregation in *late4-2* x *late3-1* - F_2 generation. In the quest of identifying potential *late4-2 late3-1* double mutants genotypically and then characterize them phenotypically, seeds from six predicted *late3-1* segregants flowering at around node 30 in F_2 were selected for sowing in F_3. This sowing comprised of 96 individuals was done in Autum (May) 2017. HRM primers used for the genotyping *late4-2* SNP have been mentioned in table A1.1.

As part of genetic interaction studies of *late3* and *late4* with *lf* and *det,* plants from already available *late3-2 lf* x *lf det* and *lf det* x *lf late4-1* populations (Weller et al., unpublished data) were grown in F_4 and F_3 generations respectively during Winter (July) 2015 to phenotypically identify putative *lf late4-1*, *lf det late3-2* and *lf det late4-1* mutants. This initially detected segregants were grown again in autumn (May) 2017 in order to confirm their phenotype and collect data for different traits. Seeds for putative *lf late3-2* mutants were provided by Jim Weller and Jackie van der Schoor. Similarly, in order to study genetic interaction of *late3* and *late4* with *sn*, plants from already generated (Weller et al., unpublished data) *sn-4* x *late3-2*, *sn-4* x *late4-1* and *sn-4* x *late4-2* populations were grown at first in Winter (September) 2016 and relevant segregant families were grown again in the next generation in Winter (July) 2017 for phenotypic characterization of the *sn-4 late4-1* double mutants only.

Phenotypic traits such as node of flower initiation (NFI), reproductive node (RN), total nodes (TN) that are used in this chapter for phenotypic characterization of WT, single, double and triple mutants have been mentioned in chapter 2. A point to note here is that phenotypic data from NGB5839, *late3-2* and *late4-1* grown in autum (May) 2017 that were used for chapter 3

have been re-used in this chapter for genetic interaction studies with *lf* and *det* as these plants were sown on the same day. Similarly, phenotypic data from NGB5839, *late4-2* and *late3-1* genotypes grown at the same time were used for genetic interaction studies between *late3* and *late4* mutants in this chapter.

6.3.2 Yeast two hybrid assay

Full length coding sequences of *PsCDK8* and *PsCYCC1* were amplified from NGB5839 genotype using Phusion high fiedelity taq polymerase system (New England Biolab). Primers used in this case have been mentioned in table A1.1. In-frame cloning of these cDNA sequences were then performed using pCR8/GW/TOPO entry vector (Invitrogen). In order to prepare bait and prey vectors for both the genes, inserts from the entry vector were then transferred into pDEST32 (bait, Invitrogen) and pDEST22 (prey, Invitrogen) destination vectors via Gateway LR recombination reaction accroding to the manufacturer's protocol of ProQuest™ Two-Hybrid System Gateway technology (Invitrogen). Primers used for checking in-frame orientation of the inserts in the destination vectors have been mentioned in table A1.1. Bait and prey vectors were prepared for both the genes.

Afterwards, haploid yeast strains PJ694 α and PJ694 A were used for transformation of bait and prey vectors respectively (de Folter and Immink 2011). PJ694α and PJ694A yeasts carrying the respective vectors were grown in synthetic (SC) complete medium lacking leucine (SC-L) and tryptophan (SC-W) respectively. Transformation was also performed using empty bait and prey vectors which were used for checking self-activation of these vectors. Transformed haploid yeasts were then allowed to mate onto synthetic medium lacking leucine and tryptophan (SC-L-W) in order to obtain diploid yeasts carrying both the vectors. Mating was peformed in both direction for each interactions. For checking interactions, diploid yeast colonies from each interactions were then mixed with 100 µl of sterile distilled water followed by 1:10 dilution in order to reduce background growth. This diluted yeast cells were then spotted in 10 µl onto selective medium lacking leucine, tryptophan, histidine (SC-L-W-H) and carrying 10 mM 3-amino-1,2,4-triazole (3AT). From each interactions, diploid yeast colonies were also spotted onto SC-L-W and SC-complete medium for comparison of growth with SC-L-W-H+3AT medium. Two independent colonies were used for each interactions. In addition, various strong and negative interactions controls were used which were created previously (Sussmilch 2014) using the constructs provided by ProQuest™ Two-Hybrid System. These

included strong positive interaction control pEXP32/Krev1 + pEXP22/RalGDS-WT; weak positive inteaction control pEXP32/Krev1 + pEXP22/RalGDS-m1 and negative inteaction control pEXP32/Krev1 + pEXP22/RalGDS-m2. Two colonies from each of these positive and negative controls were used for the interaction checking step. Moreover, duplicate colonies of diploid yeast carrying bait and empty prey vectors as well as empty bait and prey plasmids were used for checking autoactivation of the bait and prey plasmids. Likewise, two diploid colonies containing empty bait and prey plasmids were added as negative autoactivation control. Furthermore, two diploid colonies carrying no plasmids were used as negative transformation controls. Selective and non-selective medium carrying the different kinds of interactions were incubated at 30° C for four days.

6.3.3 Analysis of gene expression

Plants were grown in controlled growth chamber at 20° C (16 hour photoperiod) condition for the gene expression studies. Leaflets from penultimate fully expanded leaf from two independent plants were collected at 14, 28, 42, 56, 70 and 84 days after sowing (DAS). Samples from three such replications were harvested. Similarly, dissected stem apices of 2 mm^2 devoid of stem and sorrounding leaf tissues were collected as part of apex tissue from two independent plants for each of the three replications. Collected tissue samples were immediately placed in liquid nitrogen. RNA extraction was carried out for all three replications with the pooled leaf tissue from two plants. Similar process was followed for apex tissue as well. Protocol for total RNA extraction and cDNA synthesis (both RT+ and RT-) has been mentioned in chapter 2.

cDNA samples from both RT+ and RT- reactions of all three genotypes along with standard for a pea specific consititutive gene called *Transcription factor IIa* (*TFIIa*) and non-template control (NTC) were used for Q-RT-PCR run in order to determine the Ct values of the unknown samples. RT- samples were used to check contamination of the extracted total RNA samples with any significant level of gDNA.Two technical replications were used during the Q-RT-PCR run by using SYBR green chemistry (Sensifast, Bioline) in the following way: 5 µl Sensi fast, 0.6 µl forward primer, 0.6 µl reverse primer, 2 µl first strand cDNA and 1.8 µl sterile distilled water (SDW). In the NTC reaction, 2 µl of SDW were used instead of cDNA. Samples for Q-RT-PCR run were pipetted using CAS robotics 4 version 4.9.8 (1.6.61) software from a Corbettt Robotics CAS-1200™ pipetting robot (Corbett research, Australia). After pipetting, samples

were run in a Rotor gene Q (Qiagen) machine with the following temperature profile: 95° C-*5 minutes, 50x (*95° C- 10 seconds -> ~60° C annealing temperature-40 seconds), melt: 60-90° C with temperature rising gradually by 1° C. Similar run using primers for *FTa1*, *FTb2*, *FTc*, *PIM*, *UNI*, *VEG1*, *VEG2*, *LF* and *DET* genes were performed. Samples from all three biological replications were used for both the constitutive gene as well as aforementioned genes of interest. Relative expression of the genes of interest across different time points in three genotypes was calculated by normalizing the data against the constitutive gene *TFIIa*. Primers for the constitutive gene and genes of interest have been mentioned in table A1.1.

6.3.4 Photoperiod and vernalization experiment

For the vernalization experiment, plants from respective genotypes were kept at 4° C for 28 days after sowing, then one set of plants from each genotypes were placed in the phytotron of the School of Natural Sciences, University of Tasmania in LD (16 hour photoperiod) condition at 20° C. Another set of plants from those genotypes were placed in the phytotron under SD condition (8 hour photoperiod) at 20° C. In contrast, non-vernalized control plants from each genotypes were placed under 20° C condition at the top phytotron 21 days after launching the vernalized experiment, and they were kept under this situation for a week until they germinate. In this way, the non-vernalized plants were allowed to achieve more or less similar level of growth relative to the vernalized plants which were already kept at 4° C for 21 days. After one week, these non-vernalized plants were also splited into two groups and placed in appropriate trucks in the top phytotrons in order to continue their growth under LD and SD conditions. Data on node of flowering initiation (NFI) were collected from both the vernalized and non-vernalized plants to analyze the responsiveness of *late3* and *late4* mutations to vernalization. Data from non-vernalized control plants were also used for understanding whether *late3* and *late4* mutants regulate response to photoperiod.

6.3.5 Protein-protein interaction network for AtCDK8 and AtCYCC1

Interaction and mode of action within the components of the *A. thaliana* CDK8 module was created by using accession numbers of AtCDK8 (AT5G63610), AtCYCC1.1 (AT5G48640), AtCYCC1.2 (AT5G48630), AtMED12 (AT4G00450) and AtMED13 (AT1G55325) in STRING v10.5 database (https://string-db.org/). This was done by choosing “confidence” and “molecular action” respectively as the meaning of network edges at medium confidence interaction score (0.400) and “experiments, co-expression” as active interaction sources. In addition, the

generated network was grouped into two clusters by selecting 2 under "kmeans clustering". Larger network for AtCDK8 (AT5G63610), AtCYCC1.1 (AT5G48640) and AtCYCC1.2 (AT5G48630) was created using highest confidence interaction score (0.900), "experiments" as active interaction sources and "confidence" as the meaning of network edges and 2 for "kmeans clustering". Besides, "query proteins only" were chosen in the 1st shell and "no more than 50 interactors" in 2nd shell under the option maximum number of interactors to show in order to minimize the search results.

6.3.6 Predicted function of MtCDK8 and MtCYCC1

Predicted function of *MtCDK8* (Medtr3g096960) and *MtCYCC1* (Medtr7g055650) were developed using respective accession numbers in AraNET v2 co-functional gene network (http://www.inetbio.org/aranet/search.php) and choosing infer function from network neighbours. Methods that were used for this inference were: IMP- inferred from mutant phenotype, IGI-inferred from genetic interaction and IPI-inferred from physical interaction. This was done in order to limit generation of outcome based on known results of previous studies.

6.3.7 Predicted TF binding in CDK8 and CYCC1 gene of M. truncatula and pea

In order to determine what type of transcription factor might be regulating the function of *MtCDK8* and *MtCYCC1* in *M. trunctula*, -2000 bp to +1 bp of translation intiation site of both the genes (appendix sequence A3.1, A3.2) were collected from the *M. truncatula* Mt4.0v1 database in order to determine the potential promoter sequence where the TFs could bind. This was done because the 5'UTR of these genes are not yet annotated in the *M. truncatula* Mt4.0v1 database. However, based on RNA sequencing analysis of *PsCDK8* (previous chapter) and *PsCYCC1* transcript sequence given in pea RNA seq gene atlas database (http://bios.dijon.inra.fr/FATAL/cgi/PsUniLowCopy.cgi), it was found that the 5'UTR of these two genes in pea are 117 and 147 bp respectively. Therefore, *M. truncatula* was considered to have more or less similar length of 5'UTR for *MtCDK8* and *MtCYCC1* genes. The entire aforementioned 2 Kb region upstream of translation initiation site of *MtCDK8* and *MtCYCC1* was used to identify predictive TFs for both the genes using the binding site prediction tool (statistical threshold of 1e-5) of PlantTFDB v4.0 database (http://plantregmap.cbi.pku.edu.cn/). In this cases, only the TFs that were predicted to bind beyond the inferred 5'UTR of these two genes were considered for further analysis. Pea and

A. thaliana orthologues of these predicted *M. truncatula* TFs were identified by blastP search using respective *Medicago* protein sequences against the Pea RNA seq gene atlas (http://bios.dijon.inra.fr/FATAL/cgi/PsUniLowCopy.cgi) and Phytozyme 12 (https://phytozome.jgi.doe.gov/pz/portal.html#) databases respectively.

6.4 Results

6.4.1 LATE3 and LATE4 are most likely to function in the same regulatory pathway as unveiled by genetic interaction studies

Since *late3* and *late4* mutants showed similarities in various traits as mentioned in chapter 3, therefore the initial study carried out by Weller et al. (unpublished) prior to the commencement of the current project led to hypothesis that *LATE3* and *LATE4* act in the same regulatory pathway. This idea was strengthened by observation of non-additive effect of the *late3-1* and *late4-2* mutant phenotype among the F_2 segregants of a cross between these two mutants where a segregation ratio of 9 WT : 7 mutant was obtained (Weller et al., unpublished). Besides, identification of *LATE3* and *LATE4* as *PsCDK8* and *PsCYCC1* which are components of CDK8 module of the pea mediator complex further contributed towards the aforementioned hypothesis. In order to verify this further, seeds from six putative *late3-1* segregants (16 seeds/family) from initially crossed and already available *late4-2* x *late3-1*-F_2 population that flowered at around node 30 were sown in the F_3 generation with the anticipation of obtaining potential *late4-2 late3-1* double mutants.

The F_3 segregants were genotyped for the mutation in *late4-2* (C->T) so that potential double mutants could be identified. Out of the six F_2 families (Figure 6.1), five of those (family 1-5) showed segregation in F_3 which means the putative F_2 *late3-1* segregants were heterozygous. The remaining family (family 6) did not show segregation indicating that the respective putative F_2 *late3-1* segregant was homozygous. However, three out of these five families showed segregation for *late4-2 late1* mutation out of 16 seeds sowed for each family. A point to mention here that 24 out of the 96 seeds did not germinate properly. Nevertheless, a total of 47 homozygous *late3-1*, 14 heterozygous *late3-1* and 11 *late4-2 late3-1* segregants were detected from all the six families (Figure 6.1).

Phenotyping of the identified double mutants was then carried out for traits such as NFI, leaf area, petiole length and proximal rachis length (Figure 6.2 A-E). There was significant variation between WT and the mutants in all these traits ($p < 0.05$). Besides, no additive effect of the single mutant phenotypes were observed among the double mutants providing hints that *LATE3* and *LATE4* genes depend on each other for their function. To put it another way, these results suggest that mutation in one of these genes probably affects function of the other

gene, i.e., *LATE3* and *LATE4* are likely to act in a complementary way in order to mediate the same step of regulatory pathway that drives floral initiation as well as leaf elongation and expansion.

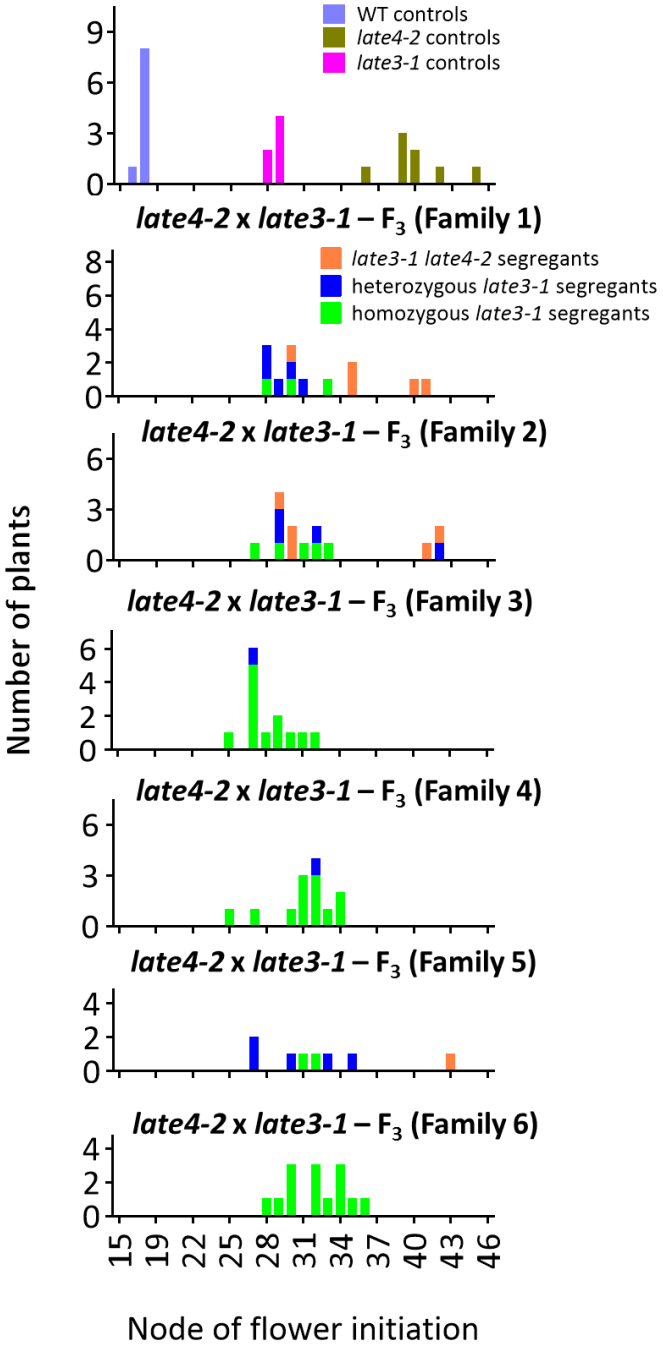


Figure 6.1. Segregation at *LATE4* locus among *late4-2* X *late3-1* progenies grown under LD condition during autumn (May) 2017. Control plants of relevant genotypes comprised of 6-9 plants (top panel). Segregation of *late3-1 late4-2* among six different families in F_3 generation grown from putative *late3-1* segregants of *late4-2* X *late3-1* – F_2 generation.

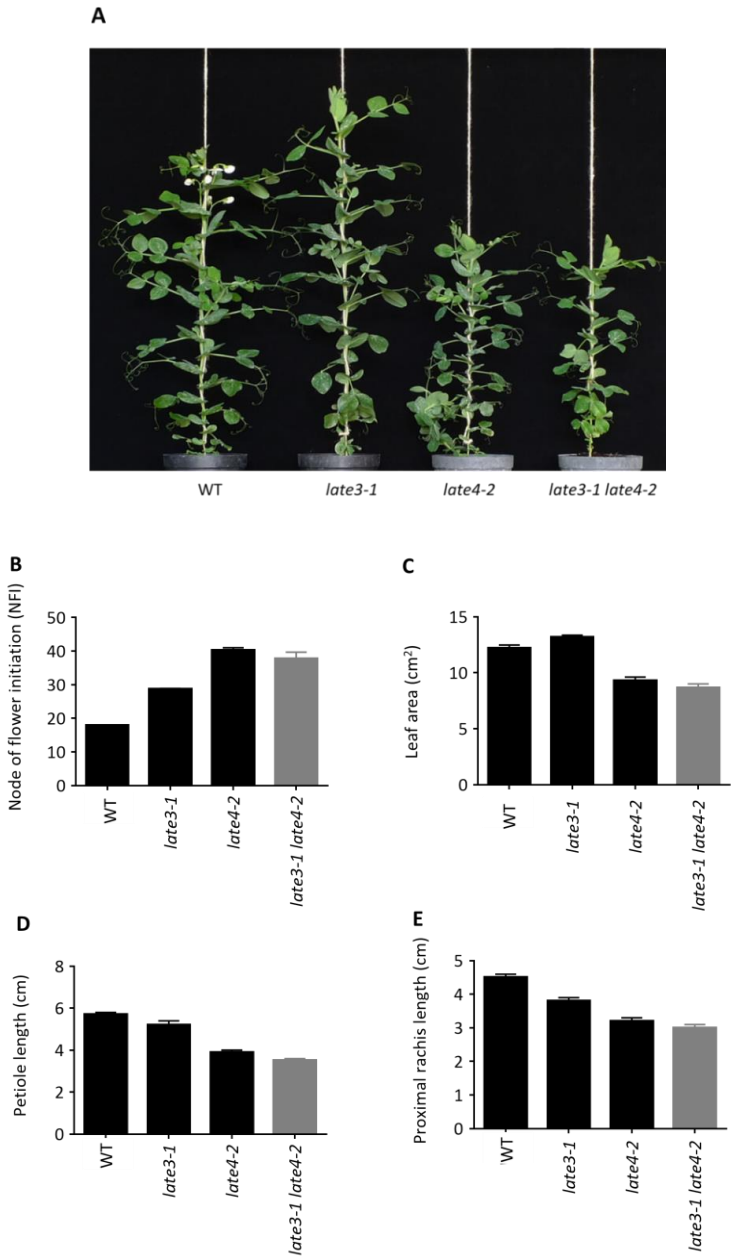
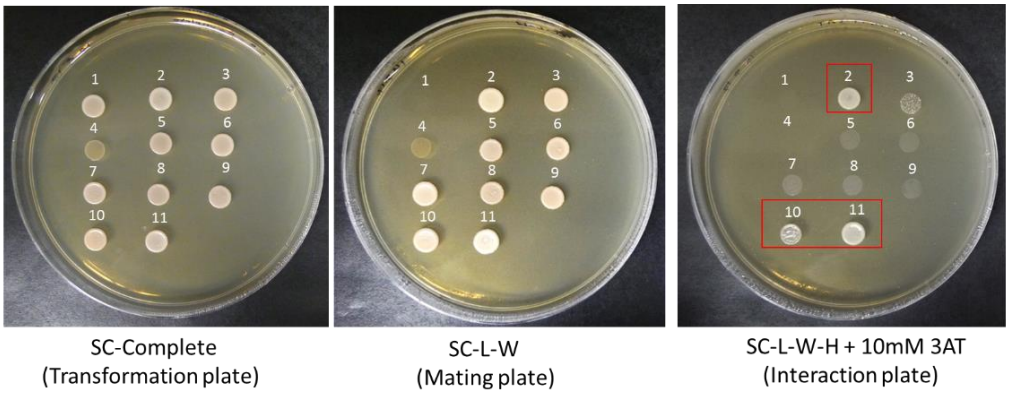


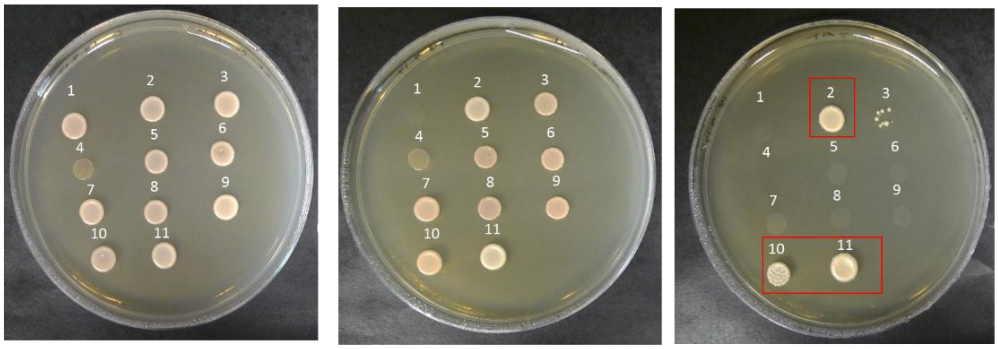
Figure 6.2. Comparison of flowering in *late3-1*, *late4-2* and *late3-1 late4-2* mutants grown under LD condition during autumn (May) 2017. **(A)** Representative 75 day old plants. **(B)** Node of flower initiation (NFI) in single and double mutants. **(C)** Leaf area which were calculated for fully expanded single leaflet of 10th leaf using imageJ software. **(D)** Petiole length which were measured for the 10th leaf. **(E)** Proximal rachis length were measured for 10th leaf. Data for each of the traits depicted in **B-E** represents mean ± SE for n = 6-10 plants.

6.4.2 Yeast two hybrid assay unravels that LATE3 and LATE4 possess strong physical interaction

Previously it was shown that CDK8 and CYCC1 proteins interact physically in *A. thaliana* (Zhu et al. 2014). Therefore, the current study hypothesized that the PsCDK8 and PsCYCC1 proteins

might be acting in the same regulatory step (as shown by genetic interaction) through direct physical interaction. Hence, investigation using yeast two hybrid assay was carried out using full length coding sequences of *PsCDK8* and *PsCYCC1* genes from the NGB5839 (WT) genotype. As part of that, different interactions including relevant positive and negative controls were checked in order to precisely understand the existence and nature of any potential physical interaction between the two proteins of interest (Figure 6.3).





Mating number	Yeast strain PJ694 α (LEU2/bait Plasmid)	Yesast strain PJ694 A (TRP1/Prey Plasmid)	Purpose	Results		
				SC-complete	SC-L-W	SC-L-W-H+3AT
1	(no plasmid)	(no plasmid)	Negative transformation control	++++	-	-
2	(pEXP32/Krev1)	(pEXP22/RPJ694 AlGDS-wt)	Strong positive interaction control	++++	++++	++++
3	(pEXP32/Krev1)	(pEXP22/RPJ694 AlGDS-m1)	Weak positive interaction control	++++	++++	+
4	(pEXP32/Krev1)	(pEXP22/RPJ694 AlGDS-m2)	Negative interaction control	+++	+++	-
5	(empty pDEST32)	(empty pDEST22)	Negative Self activation control	++++	++++	-
6	(PsCDK8_WT+pDEST32)	(empty pDEST22)	Test of self- activation bait	++++	++++	-
7	(PsCYCC1_WT+pDEST32)	(empty pDEST22)	Test of self- activation bait	++++	++++	-
8	(empty pDEST32)	(PsCDK8_WT+pDEST22)	Test of self- activation prey	++++	++++	-
9	(empty pDEST32)	(PsCYCC1_WT+pDEST22)	Test of self- activation prey	++++	++++	-
10	(PsCDK8_WT+pDEST32)	(PsCYCC1_WT+pDEST22)	Bait-specific positive interaction	++++	++++	++++
11	(PsCYCC1_WT+pDEST32)	(PsCDK8_WT+pDEST22)	Bait-specific positive interaction	++++	++++	++++

Figure 6.3. Yeast two hybrid analysis for interaction between pea Cyclin dependent kinase 8 (PsCDK8) and Cyclin C (PsCYCC1) proteins from NGB5839 genotype. Diploid yeast cells derived via mating of haploid yeast strains PJ694 α and PJ694 A carrying different bait and prey plasmids shown in the table were used for checking interaction. For each interactions, two independent mated colonies (upper panel-colony 1, lower panel- colony 2) were used which were grown in selective interaction specific (SC-L-W-H +10 mM 3AT), selective mating specific (SC-L-W) and non-selective trnasformation specific (SC-complete) medium and incubated at 30 ° C for four days. All these interactions and obtained results are mentioned in the given table where number of + signs is proportional to growth of the colony and – sign indicates no growth. Key interactions are highlighted in red color.

As was observed, diploid mated colonies carrying PsCDK8 and PsCYCC1 insert specific bait and prey plasmids (mating number 10, 11) showed growth much more than the weak positive (+) interaction control (mating number 3) in interaction specific selective medium (SC-L-W-H+10 mM 3AT) and their growth pattern were nearly similar to the strong positive interaction (+) control (mating number 2). These results suggest that PsCDK8 and PsCYCC1 proteins have very strong physical interaction. Given the complementary nature of genetic interaction between *LATE3* and *LATE4* as revealed in section 6.4.1, it is highly likely that such kind of direct physical interaction is crucial for proper functioning of these two genes.

6.4.3 LATE3 and LATE4 regulate photoperiod dependent flowering

Photoperiod is known to play important role in regulating flowering process in plants (Andres and Coupland 2012), therefore the effect of mutation in *LATE3* and *LATE4* genes on photoperiod mediated flowering in pea was checked by growing different mutant alleles under long day (LD, 16 hour photoperiod) and short day (SD, 8 hour photoperiod) conditions. As was observed, both WT and the weaker mutant allele, *late3-1* flowered significantly later under SD in comparison to LD (Figure 6.4). In contrast, the stronger mutant alleles from both the mutations, i.e., *late3-2*, *late4-1* and *late4-2* showed no significant difference in initiation of flowering between long and short day conditions. These results implicate that *LATE3* and *LATE4* genes are highly likely to act in the photoperiod regulated pathway which is responsible for commencement of flowering in pea.

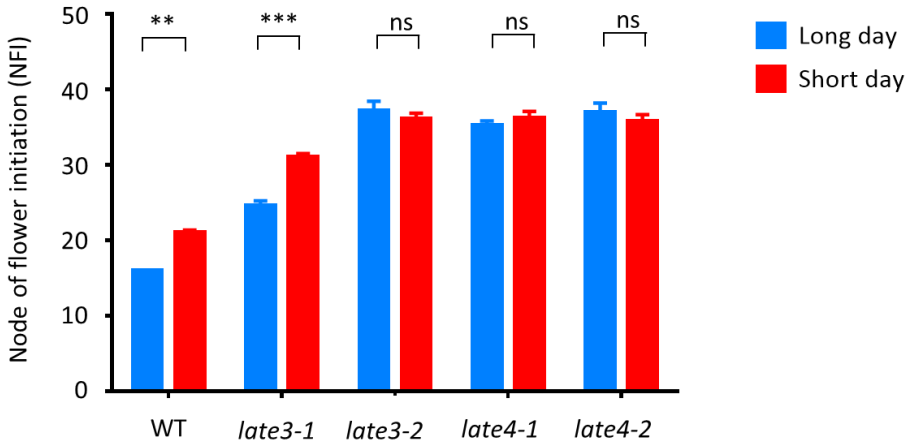


Figure 6.4. Effect of photoperiod on flowering initiation in WT (NGB5839), *late3* and *late4* mutants. Data represents mean ± SE for n = 6-8 plants. Welch t-test (two tailed) with 95% confidence interval was performed for determining statistical significance of means of long day and short day samples. p-value significance level ***≤ 0.001, ** ≤0.01, *≤ 0.05, ns > 0.05.

6.4.4 LATE3 and LATE4 are not involved in vernalization mediated flowering

Pea is a vernalization responsive plant and it was shown previously that vernalization treatment accelerates flowering in cv. Unica and Zelka grown under LD (16 hour photoperiod) (Highkin 1956) and cv. Virtus grown under SD (12 hour photoperiod) (Beveridge and Murfet 1996). In order to reveal the effect of *late3* and *late4* mutations in the vernalization mediated flowering pathway both these mutants were grown under LD and SD conditions after providing certain period of cold treatment. It was observered that under LD, WT as well as *late3-1*, *late3-2*, *late4-1* and *late4-2* mutants showed no significant difference in flowering initiation between vernalized and non-vernalized plants (Figure 6.5). In contrast, vernalized WT plants flowered significantly earlier under SD conditions in comparison to non-vernalized controls. Similar scenario was also observed for all the *late3* and *late4* mutants. Taken together, these results provide hint that mutation in *LATE3* and *LATE4* genes do not affect vernalization mediated flowering under SD conditions.

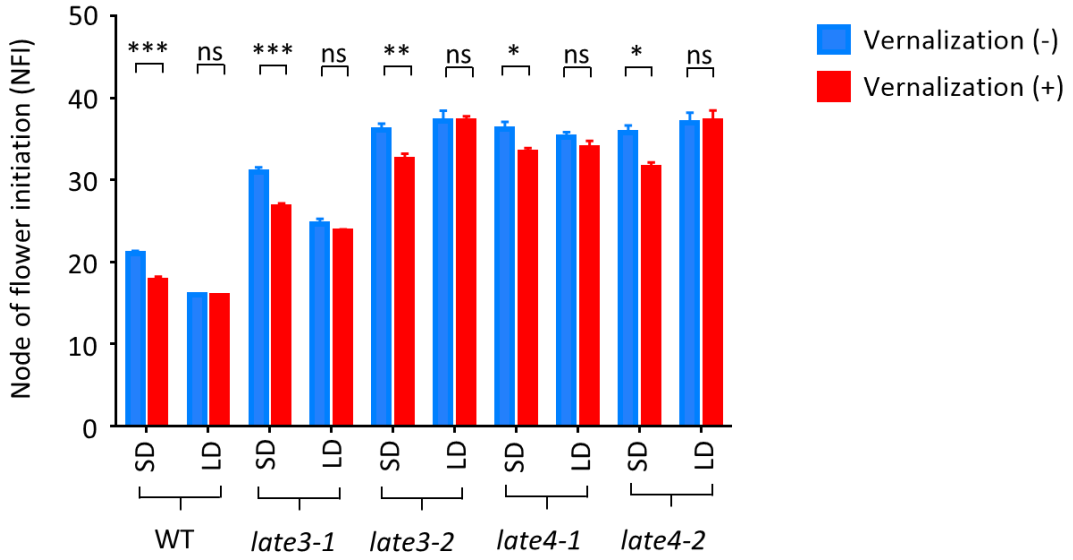


Figure 6.5. Effect of vernalization on flowering initiation in WT (NGB5839), *late3* and *late4* mutants grown under short day and long day conditions. Data represents mean ± SE for n = 6-8 plants and data from figure 6.4 have been re-plotted in this figure for non-vernalized plants of all genotypes. Welch t-test (two tailed) with 95% confidence interval was performed for determining statistical significance of means of vernalized and non-vernalized samples. p-value significance level ***≤ 0.001, ** ≤0.01, *≤ 0.05, ns > 0.05.

6.4.5 LATE3 and LATE4 are likely to promote flowering initiation and morphogenesis in pea through regulation of expression of known key pea genes

Due to the delay in flowering time as well as aberration in various developmental traits in *late3* and *late4* mutants (see chapter 3), it was assumed that the underlying genes are integral elements of pea flowering and developmental regulatory pathway and they broadly act in a positive manner for these processes. In order to understand how *LATE3* and *LATE4* genes function in mediating flowering in pea, regulatory interaction of these genes with already known key pea flowering genes such as *FTa1*, *FTb2*, *FTc*, *PIM*, *VEG1*, *VEG2*, *UNI*, *LF* and *DET* was carried out in the current project.

As was observed, expression of floral meristem identity gene, *PIM*/*PsAP1* rose dramatically from 28 days after sowing (DAS) to 42 DAS in apex of WT during the time of onset of flowering. In contrast, induction of *PIM* was delayed by nearly five weeks in *late3-1* mutants and no expression of *late4-1* was observed during the entire time period (Figure 6.6). The *late4-1* mutants grown for this experiment in the growth chamber did not flower indicating that some difference in conditions (e.g. light or temperature) may have played role in this regard. Expression pattern of another pea flower meristem identity gene *UNI*/*PsLFY* was similar to that of *PIM* in the apex of WT, however it was expressed in low level (Figure 6.6). Induction of *UNI* in *late3-1 and late4-1* occurred four and five weeks later respectively compared to WT. Expression of both these genes in the WT was consistent with previous similar studies (Hecht et al. 2011; Sussmilch et al. 2015). Notably elevated level of expression for I_2M specific gene *VEG1*/*PsFULc* in apex was observed in the WT compared to the *late3-1* and *late4-1* mutants at 28 DAS and this pattern prevailed until the emergence of flowering in the WT (Figure 6.6) in agreement to previous findings (Berbel et al. 2012; Hecht et al. 2011; Sussmilch et al. 2015). Nearly five weeks of delay in the induction of this gene was also observed in *late3-1* mutant whereas expression in *late4-1* was very low during the entire period. These results indicate that *LATE3* and *LATE4* regulate the expression of *PIM*, *UNI* and *VEG1* and thereby probably act in maintaining floral and I_2M specificity.

As far as I_1M specific *DET*/*PsTFL1a* goes, expression of this gene was very low in all three genotypes in the apex (Figure 6.6) which is in line with the fact that its expression only starts to rise after floral initiation (Sussmilch et al. 2015). Even though, the level of expression was not so high, the aforementioned pattern of *DET* expression was observed for both WT and

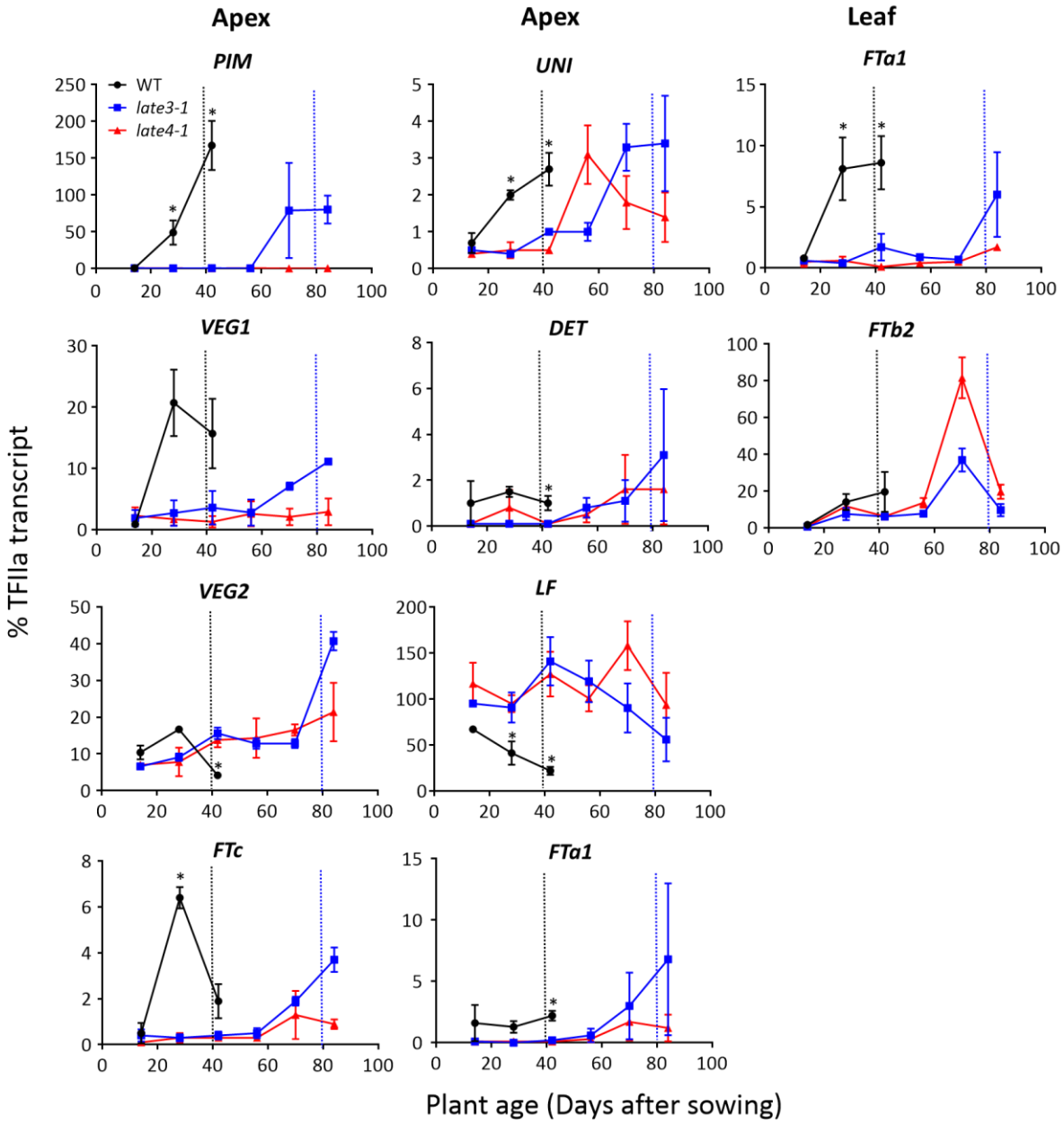


Figure 6.6. Comparative expression of key pea flowering genes in wild type NGB5839 (black circle), *late3-1* (blue square) and *late4-1* (red triangle) mutants in apex and penultimate fully expanded leaf grown under LD condition. Data have been normalized to transcript level of a pea housekeeping gene called *Transcription Factor IIa* (*TFIIa*) and represent mean ± SE for n= 3 biological replications each consisted of pooled material from 2 independent plants. Black and blue dashed line indicates initiation of flowering in WT and *late3-1* mutant respectively whereas floral initiation did not occur in *late4-1* mutants even after 154 days when plants from these genotype were discarded. One-way ANOVA followed by Dunnett test was performed to determine the level of significance ($p \leq 0.05$) between the mean of wild-type (WT) and mutant genotypes.

late3-1 mutants. Nonetheless, these results implicate that *LATE3* and *LATE4* do not have major impact on the expression of *DET* within the included time period. Expression of VM

specific gene *LF*/*PsTFL1c* in the apex was distinctly higher in the *late3-1* and *late4-1* mutants compared to the WT until 42 DAS (Figure 6.6) indicating that *LATE3* and *LATE4* act negatively upon this gene. The expression of *LF* fell off persistently from 42 DAS to 84 DAS in the *late3-1* mutant when flowering initiation occurred in this mutant which is probably consistent with reducing floral inhibitory signal derived by *LF* below to a certain threshold (Weller et al. 1997b). *VEG2* (*PsFD*) which is important for all three stages of pea inflorescence development showed higher expression in WT than *late3-1* and *late4-1* mutants at 28 DAS in apex which was followed by contrasting pattern, i.e., rapid decline at 42 DAS in WT and sharp rise in the *late3-1* and *late4-1* mutants (Figure 6.6). This implicates that *VEG2* expression is repressed by *LATE3* and *LATE4* during the time of floral initiation whereas it is promoted at earlier time point. Expression pattern of both *LF* and *VEG2* in WT was in line to previous findings (Ridge et al. 2016).

Among the *FT* genes, expression of *FTc* peaked in the apex of WT at 28 DAS compared to *late3-1* and *late4-1* mutants providing hints towards the incorporation of leaf derived floral signals by *FTc* at the apex in WT (Hecht et al. 2011) just before floral initiation (Figure 6.6). Induction of *FTc* in the mutants were delayed by five weeks. In contrast, expression of the mobile signal gene *FTa1* was slightly higher in apex of WT at the time of flowering compared to *late3-1* and *late4-1* mutants (Figure 6.6). However, *FTa1* expression was considerably higher in the leaf tissue of WT compared to both the mutants prior to flowering (Figure 6.6) which was consistent with the upregulation of FM identity genes *PIM* and *UNI* in apex of WT. There was five weeks of delay for the induction of *FTa1* in the leaf of *late3-1*. The results of *FTa1* and *FTc* expression in WT was also in agreement with that of previous studies (Hecht et al. 2011; Ridge et al. 2016). Lastly, expression of *FTb2* increased considerably from 14 DAS to 28 DAS in leaf tissues with no variation between the three genotypes (Figure 6.6). Nevertheless, relative expression of *FTb2* was much higher than that of *FTa1* in WT at 28 DAS which is probably consistent with previous findings stating that induction of the former occurs prior to the latter (Hecht et al. 2011; Ridge et al. 2016). Taken together, these results suggest that *LATE3* and *LATE4* play major role in regulating expression of *FTc* and *FTa1*, but not *FTb2* and thereby affect the timing of floral initiation.

6.4.6 Genetic interaction studies reveal potential upstream role for LATE3 and LATE4 over DET/PsTFL1a and LF/PsTFL1c in initiation of flowering

Since *late3* and *late4* mutants displayed late flowering phenotypes suggesting a positive role of the underlying genes in mediation of flowering in pea, so the current study focused on unveiling the genetic interaction with early flowering mutants such as *lf* and *det* where the underlying genes control flowering and determinacy respectively. To this end, seeds from *late3-2 lf* x *lf det*- F_3 population belonging to two putative homozygous *lf det late3-2* triple mutant families and eight putative families homozygous for *lf det*, but heterozygous for *LATE3* were sown to generate the F_4 generation. *lf late3-2* double mutant control plants were also grown. In the F_4 generation, 56 potential *lf det* double mutant segregants were identified which flowered at around node 8 similar to *lf* single mutants (Figure 6.7, top-middle panel). These *lf det* double mutants were identified based on their phenotypes which was early flowering typical of *lf* and determinacy characteristics of *det* mutants (Figure 6.8 A-B). Floral initiation for all the mutants in comparison to WT was significant ($p < 0.0001$). These result was similar to *lf det* mutants grown in a previous experiment (Sussmilch 2014). In addition, the *lf late3-2* double mutants were re-characterized based on their early flowering phenotype at around node 11 typical of *lf* and initial malformed flower typical of *late3-2* mutants (Figure 6.7, top panel). Moreover, three *lf det late3-2* triple mutants were detected from the *lf det* double mutant families in addition to confirming the phenotype of eight such putative homozygous triple mutants from the F_3 generation. These triple mutants exhibited early flowering at node 10 similar to *lf*, determinacy analogous to *det* and formation of malformed flowers along the main stem resembling that of *late3-2* (Figure 6.7, top-middle panel). In order to characterize the putative *lf det*, *lf late3-2* and *lf det late3-2* mutants further in detail and confirm their observed phenotypes in F_4 generation, only the potential homozygous lines were grown again in F_5 generation.

Similarly, seeds from *lf det* x *lf late4-1*- F_2 population belonging to two putative *lf det late4-1* triple mutant families, four putative families that were homozygous for *lf det*, but heterozygous for *LATE4* and four other families which were presumed homozygous for *lf late4-1,* but heterozygous for *DET* were sown to grow the F_3 generation. From this sowing, 29 *lf det* segregants were detected based on the same phenotypic criteria mentioned earlier

(Figure 6.7, top-lower panel). Besides, 13 *lf late4-1* double mutant segregants flowering at around node 10 were obtained which showed phenotypes similar to *lf late3-2* double mutants

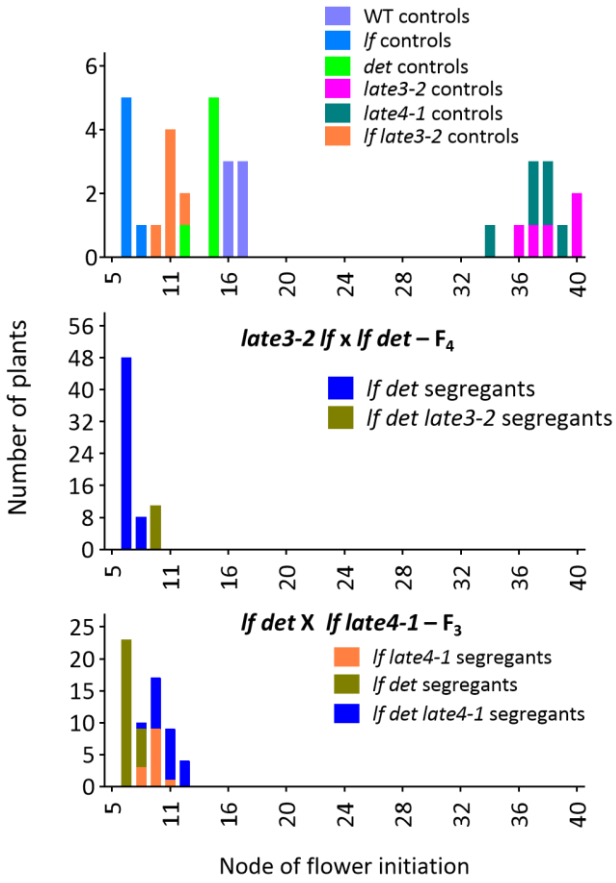


Figure 6.7. Segregation at *LATE3* and *LATE4* loci among the *late3-2 lf* x *lf det* – F_4 and *lf det* x *lf late4-1* – F_3 progenies respectively that were grown under LD condition winter (July) 2015. Control plants of relevant genotypes comprised of 5-6 plants (top panel). Putative homozygous *lf det late3-2* triple mutants and members of putative families homozygous for *lf det* (but heterozygous for *LATE3*) from *late3-2 lf* x *lf det* – F_3 were grown in F_4 generation (middle panel). Putative homozygous *lf det late4-1* triple mutants and members of putative families homozygous for *lf det* (but heterozygous for *LATE4*) and homozygous for *lf late4-1* (but heterozygous for *DET*) from *lf det* x *lf late4-1*- F_2 were grown in F_3 generation (lower panel).

mentioned earlier (Figure 6.7, top-lower panel). 13 *lf det late4-1* triple mutants were also identified in addition to confirming phenotype of eight such putative triple mutants obtained from earlier generation (Figure 6.7, top-lower panel). These potential triple mutants flowered around node 10-11 and their phenotypes were similar to that of *lf det late3-2* mutants. Putative homozygous *lf late4-1* and *lf det late4-1* mutants were grown in next generation in order to characterize them further and also to confirm their observed phenotypes from the F_3 generation.

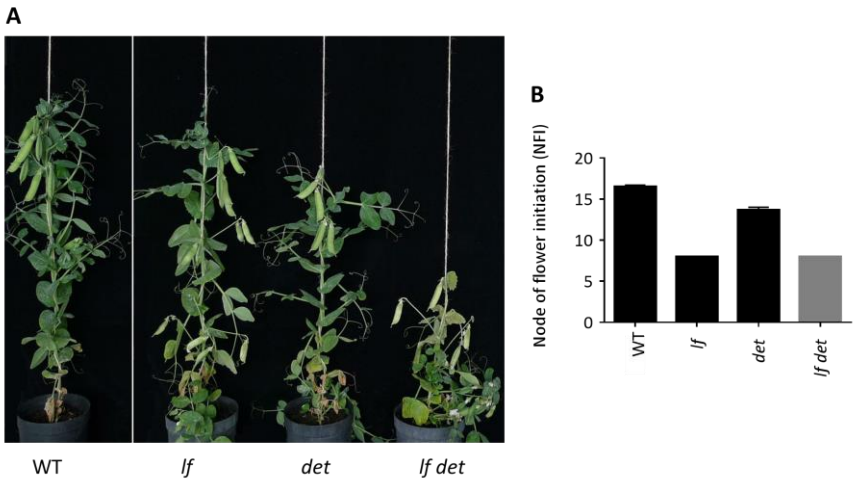


Figure 6.8. Comparative flowering in wild type (WT) NGB5839, *lf*, *det* and *lf det* mutants grown under LD during winter (July) 2015. (**A**) Representative 83 day old plants. (**B**) Node of flower initiation in each of the genotype, data represents mean ± SE for n = 6 plants.

During autumn (May) 2017, the aforementioned putative homozygous *lf det*, *lf late3-2*, *lf late4-1*, *lf det late3-2* and *lf det late4-1* mutants were grown again from selected lines along with the relevant single mutants and WT controls. As far as the putative *lf late3-2* and *lf late4-1* mutants were concerned, they flowered around node 11 and 10 respectively which confirmed this particular phenotype from the F_3 generation (Figure 6.9 A-B). These results suggest that *LF* is epistatic to *LATE3* and *LATE4* for the flower initiation phenotype and acts downstream of *LATE3* and *LATE4* gene. However, both the double mutants showed longer reproductive phase similar to the *late3-2* and *late4-1* single mutants (Figure 6.9 B-C). Both Floral initiation and reproductive phase for all the mutants in comparison to WT was statistically significant ($p < 0.0001$). Comparison of the reproductive phase with *lf* single mutants indicate that there is no genetic interaction, rather additive effect of *late3-2* and *late4-1* mutation for this particular trait. In addition, the double mutants did not show massive branching that was observed in the *late3-2* and *late4-1* single mutants (Figure 6.9 C).

The putative *lf det late3-2* and *lf det late4-1* triple mutants that were grown during autumn (May) 2017 flowered at around node 10 nearly similar to *lf* single and *lf det* double mutants again confirming the phenotype observed in the F_4 generation (Figure 6.10 A, D-E). Besides, the triple mutants had a shorter reproductive phase similar to *lf det* double mutants (Figure 6.10 D-E). There was significant variation ($p < 0.05$) for the floral initiation and reproductive phase data of all the mutants relative to WT. These results suggest upstream role for *LATE3*

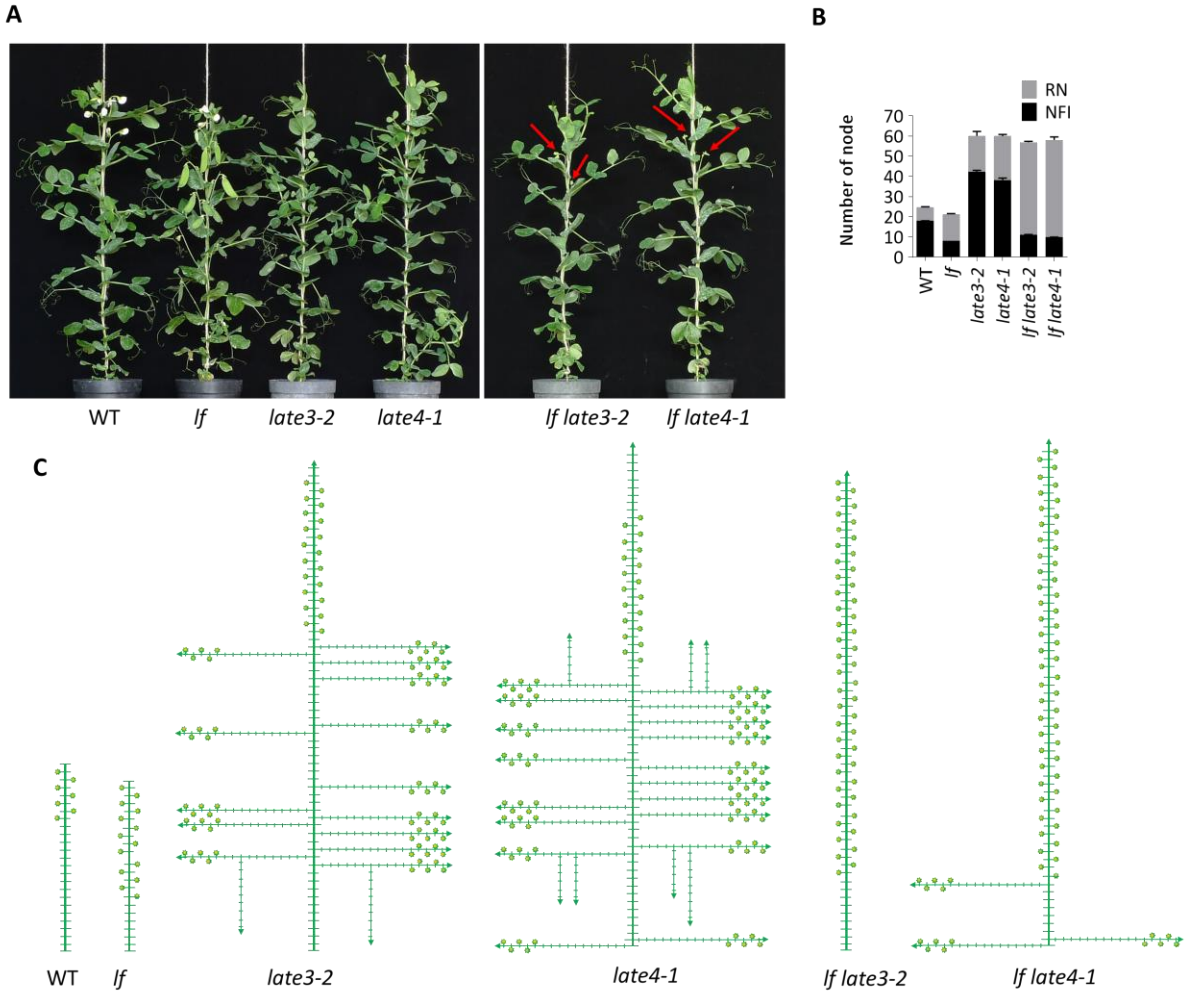


Figure 6.9. Comparative flowering in wild type (WT) NGB5839, *lf*, *late3-2*, *late4-1*, *lf late3-2* and *lf late4-1* mutants grown under LD during autumn (May) 2017. **(A)** Representative 75 day old plants, red colored arrows indicate malformed flowers. **(B)** Node of flower initiation (NFI) and reproductive node (RN) in each of the genotype, data represents mean ± SE for n = 6-10 plants. Data from figure 3.1 have been re-plotted. **(C)** Schematic diagram comparing morphological structure observed at the time of harvest (180 days after sowing) in wild type (WT) NGB5839, *lf*, *late3-2*, *late4-1*, *lf late3-2* and *lf late4-1* mutants. Long vertical line represents main stem, small horizontal lines in the main stem indicate a node where the line at the bottom is node 1, horizontal lines coming out from a node represent primary lateral branches while additional branching in presented by further vertical lines, star shaped green and yellow structure represent reproductive organ, single arrow indicates continuous growth.

and *LATE4* over the activity of *LF* and *DET* in relation to promotion of flowering and duration of reproductive phase in pea. The triple mutants generated malformed flowers in all the reproductive nodes along the main stems (Figure 6.10 B-C) and it was contrasting with *lf det* double mutants which produced normal flowers in the main stem leading to seed production (Figure 6.8 A, 6.10 E). These results indicate more severe effect of the *late3-2* and *late4-1* mutations on floral fertility in combination to *lf det* mutation. The triple mutants were also

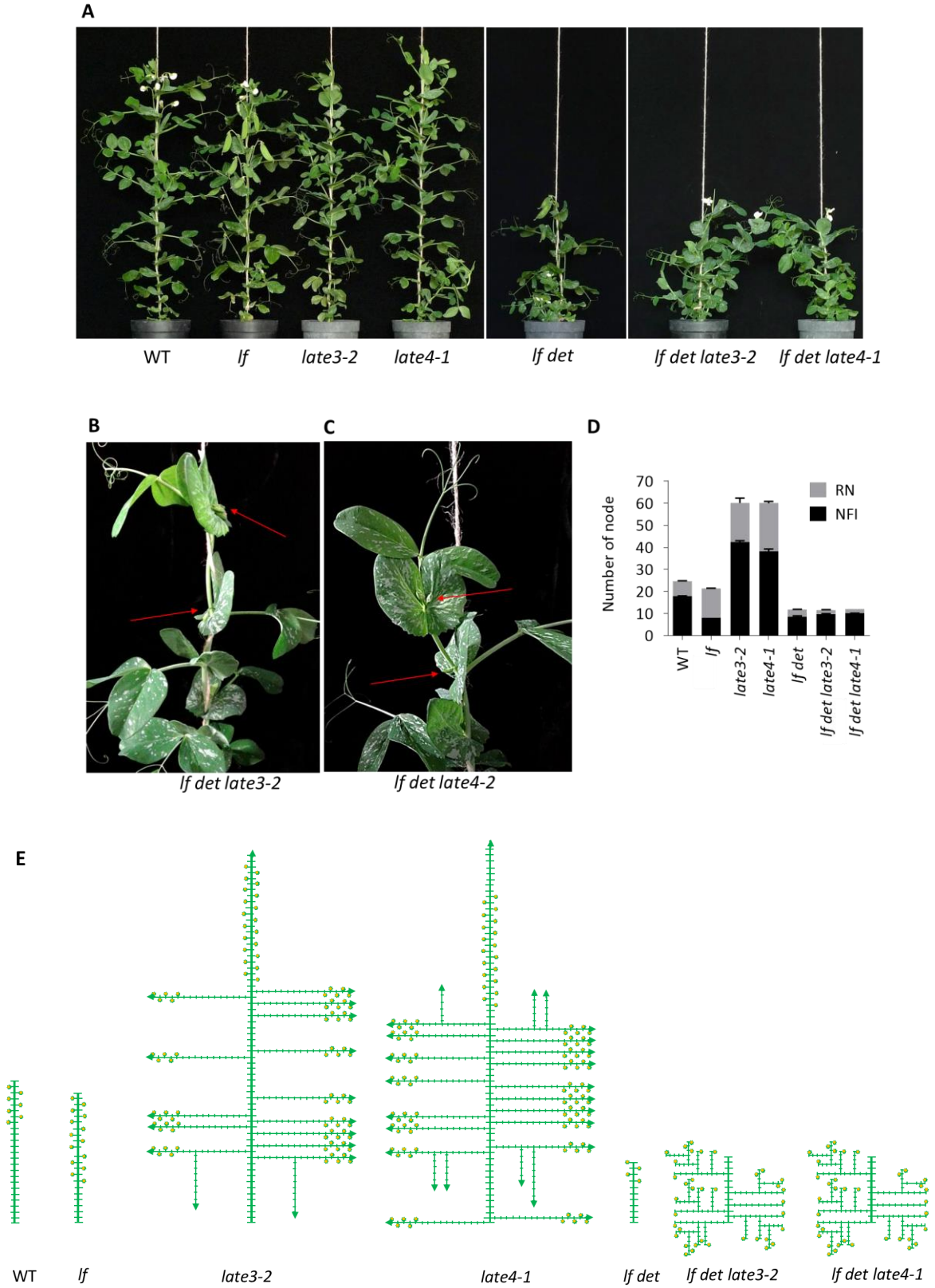


Figure 6.10. Comparative flowering in wild type (WT) NGB5839, *lf*, *late3-2*, *late4-1*, *lf det, lf det late3-2* and *lf det late4-1* mutants grown under LD during autumn (May) 2017. **(A)** Representative 75 day old plants grown under LD. **(B-C)** Malformed initial flowers in the main stem of *lf det late3* and *lf det late4* triple mutants. **(continued next page)**

Figure 6.10. (continued) (D) Node of flower initiation (NFI) and reproductive nodes (RN), data represents mean ± SE for n = 6-10 plants. Data from figure 3.1 have been re-plotted. **(E)** Schematic diagram comparing morphological structure observed at the time of harvest (180 days after sowing) in WT (NGB5839), *lf*, *late3-2*, *late4-1*, *lf det, lf det late3-2* and *lf det late4-1* mutants. Long vertical line represents main stem, small horizontal lines in the main stem indicate a node where the line at the bottom is node 1, large horizontal lines coming out from a node in the main stem represent primary lateral branches while additional branching is presented by further horizontal/vertical lines, star shaped green and yellow structure represent reproductive organ, arrow indicates continuous growth.

vastly branched with up to 4-5 orders of branching whereas the *late3-2* and *late4-1* single mutants showed mostly primary branching with occasional secondary branches (Figure 6.10 E). It could therefore be stated that the branching phenotype of *late3-2* and *late4-1* mutants also gets more acute when combined with *lf det* double mutation. Unlike *late3-2* and *late4-1* single mutants, these branches were determinate which produced flowers as well as seeds. Both *det* and *lf det* mutants are known to show determinate growth with shoot apices finish growth with a terminal flower (Figure 6.8 A) (Murfet 1989; Singer et al. 1990; Sussmilch 2014). This particular phenotype was observed for *lf det late3-2* and *lf det late4-1* triple mutants in the main stem (Figure 6.10 B-C) and flowering branches indicating that mutation in *late3* and *late4* genes does not alter the characteristic feature of *det* and *lf det* mutants for this trait, rather increases its severity.

6.4.7 Pea circadian clock gene SN/PsLUX seems to act upstream of LATE4 to repress flowering

In order to understand how *LATE3* and *LATE4* promote flowering in pea, additional genetic interaction studies was carried out in the current project by using early flowering *sn-4* mutants where the underlying gene inhibit expression of key *FT* genes (Liew et al. 2014).

To this end, *sn-4* x *late3-1* F_4 population was generated during spring (September) 2016 from four different *sn-4* segregant families of F_3 generation. All the plants from F_3 generation turned out to be homozygous as they flowered around node 12 similar to *sn-4* controls and thereby no *sn-4 late3-2* double mutant segregant was obtained in the F_4 generation (Figure 6.11 first and second panel). Besides, one seed obtained from a potential homozygous *sn-4 late3-2* double mutant in F_3 generation was also sown in the F_4 generation, but this seed did not germinate. So, seeds from this particular cross could not be used further phenotypic studies in the next generation.

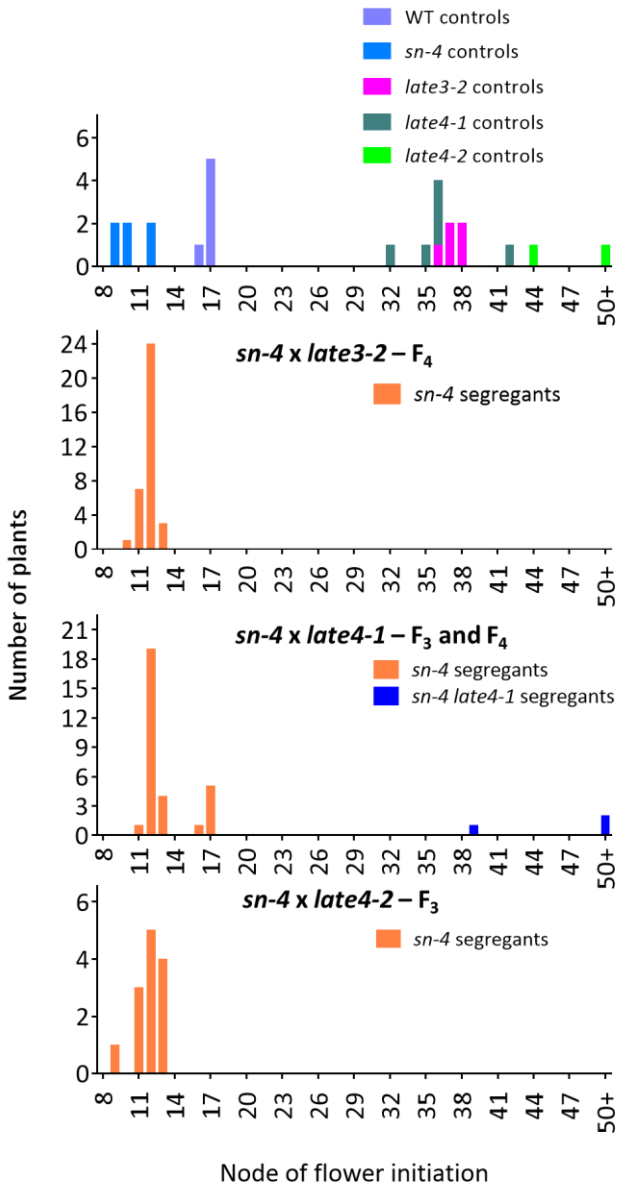


Figure 6.11. Segregation at *LATE3* and *LATE4* loci among the *sn-4* x *late3* and *sn-4* x *late4* progenies respectively grown under LD condition during spring (September) 2016. Control plants of relevant genotypes comprised of 5-10 plants (first or top panel). Siblings (*sn-4* phenotype) of *sn-4 late3-2* segregants from *sn-4* x *late3-2*-F_3 were grown in F_4 generation (second panel). Siblings (*sn-4* phenotype) of *sn-4 late4-1* segregants from *sn-4* x *late4-1* - F_2 and *sn-4* x *late4-1* – F_3 were grown in F_3 and F_4 generations respectively (third panel). Siblings (*sn-4* phenotype) of *sn-4 late4-2* segregants from *sn-4* x *late4-2* - F_2 were grown in F_3 generation (fourth or bottom panel).

During the same time in 2016, *sn-4* x *late4-2* F_3 progenies were grown from the lone *sn-4* segregant family of F_2 generation. In this case, all the plants from F_2 generation seemed to be homozygous as well since flowering initiated for all of them in F_3 at around node 12 alike *sn-4* single mutants (Figure 6.11 first and third panel). Therefore, seeds from this cross also could

not be used further for characterization of potential *sn-4 late4-2* double mutants. The third group of plants that were grown at the same period in 2016 comprised of a *sn-4* segregant family of *sn-4* x *late4-1* –F_2 population as well as six different *sn-4* segregant families of *sn-4* x *late4-1* – F_3 population. Out of these seven *sn-4* segregant families, three families showed segregation for *late4-1* phenotype in subsequent generations, thus three *sn-4 late4-1* double mutants were obtained. Two of the double mutants did not show floral initiation even after 50 nodes whereas another one flowered at node 39 similar to *late4-1* controls (Figure 6.11, first and fourth panel). None of these double mutants generated any seeds indicating that *sn-4* and *late4-1* in combination is more severe for floral fertilization and they have overlapping functions in the production of seeds. Due to this aspect, it was needed to obtain the *sn-4 late4-1* double mutants again by growing seeds of *sn-4* segregant families of the respective generations.

During winter (July) 2017, five different *sn-4* segregant families of *sn-4* x *late4-1* F_3 population and four different *sn-4* segregant families of *sn-4* x *late4-1* F_4 population were sown in next generation. As expected according to Mendelian segregation ratio, six out of the nine families turned out to be heterozygous as they showed segregation for *sn-4 late4-1* double mutation (Figure 6.12, 6.13 A-B). A total of nine *sn-4 late4-1* double mutants were obtained from these

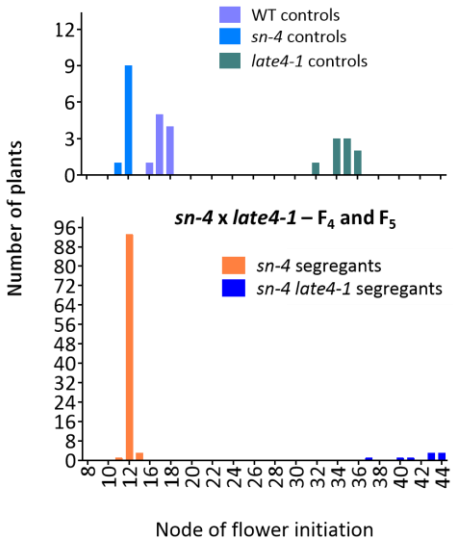


Figure 6.12. Segregation at *LATE4* loci among the *sn-4* x *late4-1* progenies respectively grown under LD condition during winter (July) 2017. Control plants of relevant genotypes comprised of 5-10 plants (upper panel). Number of *sn-4* and *sn-4 late4-1* segregants obtained from *sn-4* x *late4-1* – F_4 and F_5 generations (lower panel).

families which flowered at around node 43. There was significant difference ($p < 0.0001$) in floral initiation for all the mutants compared to WT. This result suggests that *LATE4* is epistatic to *SN* and acts downstream of *SN* in the same regulatory pathway that mediate floral initiation in pea. Besides, these double mutants did not generate any seeds similar to the scenario observed in the previous generation confirming again acute phenotype for this trait when *sn-4* and *late4-1* mutations are combined.

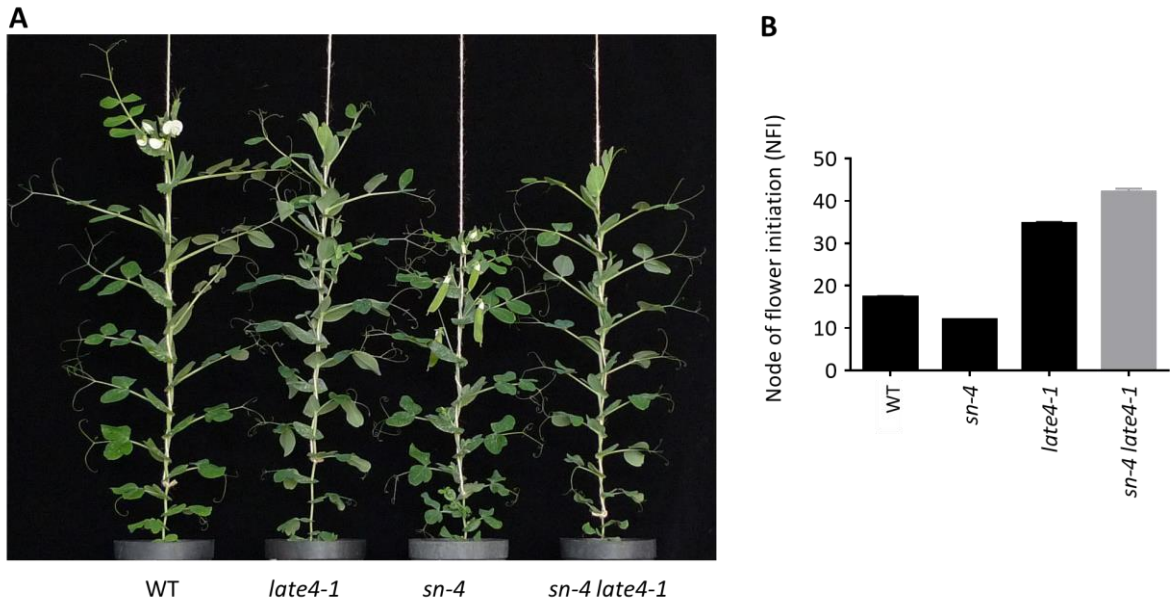


Figure 6.13. Comparative flowering in wild type (WT) NGB5839, *late4-1*, *sn-4* and *sn-4 late4-1* mutants grown under LD during winter (July) 2017. **(A)** Representative 63 day old plants. **(B)** Node of flower initiation in each of the genotype, data represents mean ± SE for n = 6-10 plants.

6.4.8 Generation of protein-protein interaction network for A. thaliana AtCDK8 and AtCYCC1 using STRING v10.5

As mentioned earlier, components of the CDK8 module of eukaryotic mediator complex have vital role in regulating transcription of various genes in the eukaryotic system (Yang et al. 2016). Some studies have been performed in animal systesm such as *Drosophila* and humans about the nature of interaction between CDK8, CYCC1, MED12 and MED13 of CDK8 module which then interacts with the large core mediator in regulating transcription (Loncle et al. 2007; Tsai et al. 2013). However, knowledge about the interaction between components of CDK8 module in plant system is limited. The current study initially aimed at addressing the interaction of various components of the pea CDK8 module through yeast two hybird (Y2H)

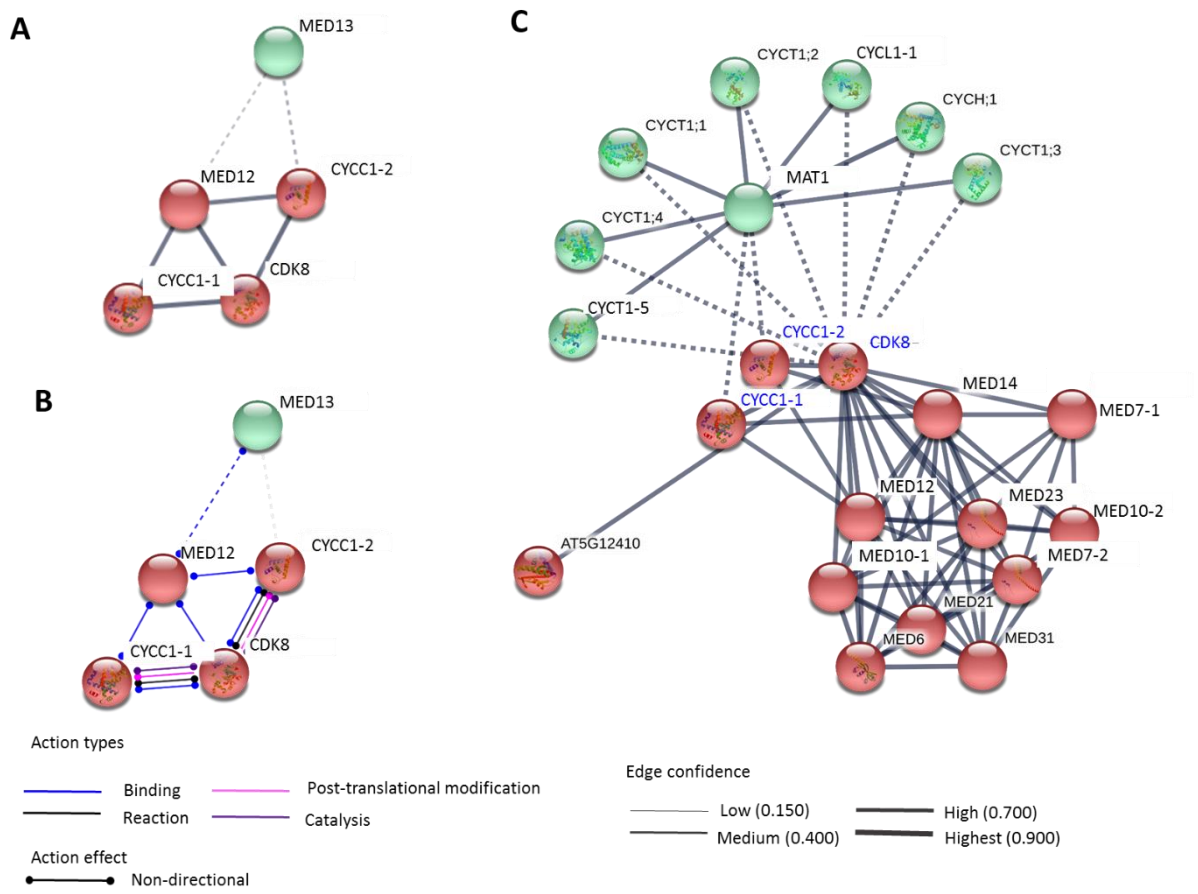


Figure 6.14. Protein-protein association network for AtCDK8, AtCYCC1.1 and AtCYCC1.2 generated using STRING V10.5 database (https://string-db.org/). **(A-B)** Interaction and molecular action between different components of the AtCDK8 module at medium confidence interaction score (0.400). **(C)** Larger interaction network for showing association with various proteins at highest confidence interaction score (0.900). Each of the circles represent single gene (node) whereas lines (edges) represent interaction between them. All three networks have been divided into two clusters (green and red color).

method. But, the *PsMED12-1*, *PsMED12-2* and *PsMED13* genes were found to have ORFs with length of approximately 6.9, 6.7 and 5.9 Kb respectively. Therefore, focus was limited to the smaller *PsCDK8* and *PsCYCC1* genes which have ORFs of nearly 1.4 and 0.75 Kb. Nevertheless, given the importance of the CDK8 module in eukaryotic transcriptional regulation, the interaction network database namely STRING v10.5 (Szklarczyk et al. 2015) was used in order to understand the predicted mode of interaction between different genes of the *A. thaliana* CDK8 using medium confidence score (0.400) and available experiemental and co-expression data (Figure 6.14 A). Proteins interacting with each other are likely to have complementary expression profiles (Fraser et al. 2004; Ge et al. 2001), therefore co-expression data was

selected in addition to experimental results for checking the interaction between the various components of AtCDK8 module. In this case, strong network edges showing predicted functional association between AtCDK8, AtCYCC1-1, AtCYCC1-2 and AtMED12 has been observed not only in medium confidence score (Figure 6.14 A), but also at highest confidence score (0.900). In contrast, interaction for AtMED13 was not observed at highest (0.900) or high (0.700) confidence score with any other component, rather only at medium (0.400) confidence score with AtMED12 and AtCYCC1-2. In addition, findings of predicted molecular action showed that AtMED12 is likely to play a central role in order to contribute towards a shared function for the AtCDK8 module (Figure 6.14 B). These results are consistent with the findings in yeast and human via electron microscopy and biochemistry to understand subunit organization, structure and interaction of CDK8 module where MED13, CDK8 and CYCC1 were proposed to be associated around centrally placed MED12 (Wang et al. 2013; Tsai et al. 2013; Jeronimo and Robert 2017).

A broader protein-protein interaction network of AtCDK8, AtCYCC1-1 and AtCYCC1-2 was also created in order to identify the potential interacting partners of the aforemtioned proteins based on experimental data at highest confidence intereaction score (0.900) (Figure 6.14 C, Table 6.1). The result generated 19 additional proteins showing different degrees of interaction with the three aforementioned mediator proteins of *A. thaliana.* As was observed, these proteins were mostly components of mediator complex, e.g. MED6, MED7, MED10, MED12, MED21, MED23, MED31 etc. belonging to one cluster (red color) and cell cycle such as CYCT1-1, CYCT1-2, CYCT1-3, CYCT1-4, CYCT1-5, CYCL1-1, CYCH1-1, MAT1 forming a separate cluster (green color). Such results were expected since *CYCC1* has known role in cell cycle and mediator complex in other systems whereas *CDK8* in mediator complex only (Sage 2004; Allen and Taatjes 2015).

Table 6.1. Names and annotated functions of of various network proteins of AtCDK8, AtCYCC1-1 and AtCYCC1-2. Query proteins are marked as blue.

Node	Accession numbers	Annotation
MED23	AT1G23230	Mediator complex component
MED10-1	AT1G26665	Mediator complex component
CYCT1-3	AT1G27630	Cell cycle component
CYCT1-1	AT1G35440	Cell cycle component
CYCL1-1	AT2G26430	Cell cycle component
MED14	AT3G04740	Mediator complex component
MED6	AT3G21350	Mediator complex component
MED12	AT4G00450	Mediator complex component
MED21	AT4G04780	Mediator complex component
CYCT1-2	AT4G19560	Cell cycle component
CYCT1-4	AT4G19600	Cell cycle component
MAT1	AT4G30820	Cell cycle component
MED7-1	AT5G03220	Mediator complex component
MED7-2	AT5G03500	Mediator complex component
AT5G12410	AT5G12410	THUMP domain-containing protein
MED31	AT5G19910	Mediator complex component
CYCH1-1	AT5G27620	Cell cycle component
MED10-2	AT5G41910	Mediator complex component
CYCT1-5	AT5G45190	Cell cycle component
CYCC1-2	AT5G48630	Mediator complex/Cell cycle component
CYCC1-1	AT5G48640	Mediator complex/Cell cycle component
CDK8	AT5G63610	Mediator complex component

6.4.9 Data from AraNET and PlantTFDB provides useful information about potential function of MtCDK8/MtCYCC1 and regulation of these genes

Since *LATE3* and *LATE4* genes have been identified as pea homologues of eukaryotic mediator complex component *CDK8* and *CYCC1* respectively, therefore it is expected that function these genes are likely to be conserved across various legumes. In order to compare the results obtained in the present study for *PsCDK8* and *PsCYCC1*, the co-functional gene network platform AraNET v2 was used for predicting the function of both these genes in *M. truncatula*, a species which is highly syntenic to pea. Results obtained by selecting 'infer functions from network neighbours' (inferred from known mutant phenotype, genetic and physical interaction) unveiled different predicted functions of *MtCDK8* and *MtCYCC1* genes (Table 6.2, 6.3). For both the genes, 30 top ranked functions were selected. These functions (GO term)

Table 6.2. Predicted functions of *MtCDK8* (Medtr3g096960) gene inferred from network neighbours developed using AraNet v2 (http://www.inetbio.org/aranet/) co-functional gene network.

Rank	Score	GO term	GO term supporters(LLS)
1	29.74	Flower development	HAT1(7.30) CDKC;2(5.32) AT5G45190(3.17) HAC12(2.40) PCAT2(2.40) HAC5(1.85)
2	25.21	Regulation of G2/M transition of mitotic cell cycle	CDKB2;2(4.78) CDKB2;1(4.71) CYCA2;4(3.70) CYCA2;2(3.70) CYCA2;3(3.54)
3	23.12	Leaf development	CDKC;1(5.46) CDKC;2(5.32) CYCT1;4(3.70) AT5G45190(3.17)
4	22.83	Carpel development	MED18(8.76) CDKC;2(5.32)
5	19.96	Response to water deprivation	CYCH;1(7.99) CHR12(3.97)
6	19.89	Regulation transition from vegetative to reproductive phase	MED18(8.76) NUC(2.38)
7	19.38	DNA endoreduplication	CDKA;1(4.05) CYCA2;1(3.87) CYCD5;1(3.87) CYCA2;3(3.54)
8	18.81	Reactive oxygen species metabolic process	CYCH;1(7.99) GPA1(2.82)
9	18.81	Regulation of stomatal movement	CYCH;1(7.99) GPA1(2.82)
10	18.69	Cell proliferation	CYCA2;1(3.87) CYCA2;4(3.70) CYCA2;2(3.70) CYCA2;3(3.54)
11	17.74	Response to light stimulus	HAT1(7.30) TIL(3.14)
12	17.53	Regulation of cell proliferation	CYCD3;3(4.78) CYCD3;2(3.40) GPA1(2.82) PRZ1(1.75)
13	17.51	Negative regulation of transcription, DNA-templated	MED18(8.76)
14	17.51	Petal development	MED18(8.76)
15	17.51	Specification of floral organ number	MED18(8.76)
16	17.51	Sepal development	MED18(8.76)
17	17.51	Regulation of vernalization response	MED18(8.76)
18	17.51	Stamen development	MED18(8.76)
19	15.99	Regulation of response to water deprivation	CYCH;1(7.99)
20	15.99	Stomatal opening	CYCH;1(7.99)
21	14.68	Embryo development ending in seed dormancy	CDKA;1(4.05) SCC2(2.82) TOP1ALPHA(1.97) CUL3A(1.80)
22	14.6	Root morphogenesis	HAT1(7.30)
23	14.27	Hormone-mediated signalling pathway	CDKB2;2(4.78) CDKB2;1(4.71)
24	14.27	Regulation of meristem structural organization	CDKB2;2(4.78) CDKB2;1(4.71)
25	13.82	Regulation of viral process	CDKC;2(5.32) AT5G45190(3.17)
26	12.95	Seed development	CYCD3;3(4.78) CYCD3;2(3.40)
27	12.19	Embryo development	QQT2(3.60) AT1G63210(2.53) eIF3F(2.46)
28	12.05	Defence response to fungus	MED21(4.02) PFT1(4.00)
29	11.09	Response to heat	CHR12(3.97) TIL(3.14)
30	10.95	Positive regulation of stomatal complex development	CYCA2;2(3.70) CYCA2;3(3.54)

Table 6.3. Predicted functions of *MtCYCC1* (Medtr7g055650) gene inferred from network neighbours developed using AraNet v2 (http://www.inetbio.org/aranet/) co-functional gene network.

Rank	Score	GO term	GO term supporters(LLS)
1	25.47	Cell proliferation	CYCA2;1(4.50) CYCA2;4(4.50) CYCA2;3(4.50) CYCA2;2(4.50) EMB2783(2.98)
2	23.5	Response to water deprivation	CYCH;1(8.90) CHR12(2.89) HSP90.2(2.81)
3	20.43	Response to heat	VTC1(4.26) HSP70-1(3.94) CHR12(2.89) Hsp81.3(2.82) HSP101(2.26)
4	20.42	DNA endoreduplication	CYCD5;1(4.50) CYCA2;1(4.50) CYCA2;3(4.50) CDKA;1(2.42)
5	20.26	Regulation of cell proliferation	CYCD3;3(4.50) CYCD3;2(4.50) CYCD3;1(4.50) MCM2(2.25)
6	19.84	Flower development	HAT1(7.61) HSP90.2(2.81) CUL4(1.80)
7	19.63	Defence response to bacterium	VTC1(4.26) HSP70-1(3.94) HSP90.1(2.83) PEN3(2.22) PMR4(2.12)
8	18.01	Seed development	CYCD3;3(4.50) CYCD3;2(4.50) CYCD3;1(4.50)
9	18	Regulation of G2/M transition of mitotic cell cycle	CYCA2;4(4.50) CYCA2;3(4.50) CYCA2;2(4.50)
10	17.8	Stomatal opening	CYCH;1(8.90)
11	17.8	Regulation of response to water deprivation	CYCH;1(8.90)
12	17.8	Reactive oxygen species metabolic process	CYCH;1(8.90)
13	17.8	Regulation of stomatal movement	CYCH;1(8.90)
14	17.28	Response to salt stress	VTC1(4.26) J3(3.07) CHR12(2.89) HSP90.2(2.81)
15	15.77	Defence response to fungus	HSP70-1(3.94) MED21(2.89) PFT1(2.88) PMR4(2.12)
16	15.51	Microsporogenesis	TAM(4.51) GLS2(2.26) GSL10(2.12) GSL8(2.12)
17	15.46	Regulation of timing of transition from vegetative to reproductive phase	MED18(5.48) NUC(4.50)
18	15.3	Endosperm development	TFIIB1(5.04) CUL3B(2.66) CUL3A(2.57)
19	15.23	Root morphogenesis	HAT1(7.61)
20	15.23	Response to light stimulus	HAT1(7.61)
21	14.45	Regulation of pollen tube growth	TFIIB1(5.04) GLS2(2.26) GSL10(2.12)
23	13.5	Positive regulation of stomatal complex development	CYCA2;3(4.50) CYCA2;2(4.50)
24	13.49	Regulation of starch metabolic process	SGR5(4.50) IDD14beta(4.50)
25	13.45	Stamen development	MED18(5.48) NUA(2.49)
26	13.39	Stomatal lineage progression	CYCD4;1(4.50) CYCD4;2(4.38)
27	13.29	Pollen development	CDKA;1(2.42) GSL10(2.12) GSL8(2.12) PMR4(2.12) GSL1(2.11)
28	12.83	Embryo development ending in seed dormancy	CUL3B(2.66) CUL3A(2.57) SKP1B(2.53) CDKA;1(2.42)
29	12.07	Photoperiodism, flowering	NUC(4.50) J3(3.07)
30	11.59	Positive regulation of flower development	J3(3.07) PFT1(2.88) CUL3A(2.57)

included mostly role in flower and leaf development, transition from vegetative to reproductive phase, floral morphogenesis, seed development, root morphogenesis and response to various abiotic and biotic factors. Most of these results related to reproduction

and development are similar to what was observed in pea in the current study providing hints that the universal function of the *CDK8* and *CYCC1* genes likely to be conserved between the two species.

Transcription factors are known to play major role in regulating expression of genes at the transcriptional level and thereby determine the function of the genes (Mitsuda and Ohme-Takagi 2009; Yang et al. 2016). A number of plant transcription factors possess a DNA-binding domain (DBD) through which they bind to the promoter or further upstream at the enhancer sequences of respective genes upon receiving various endogenous/exogenous signals. In the quest of detecting the transcription factors that are likely to regulate function of *MtCDK8* and *MtCYCC1* gene by directly binding to their promoter sequences, the PlantTFDB4.0 database was exploited which has a large repertoire of plant transcription factors mentioned earlier.

For *MtCDK8*, some of the top ranked predicted transcription factors (TFs) included those from AP2, BBR-BPC, bZIP, C3H, Dof, ERF, GRAS and MYB family (Table 6.4). The annotated role of these classes of TFs include regulation of wide range of biological processes such as flower, root, shoot, leaf, embryo, seed development, biotic and abiotic stress response, plant cell fate determination, carbohydrate and secondary metabolism. In contrast, top ranked TFs for *MtCYCC1* also included AP2, BBR-BPC, bZIP, Dof, ERF and GRAS classes. In addition, four other types of TF namely CPP, HD-ZP, MIK-MADS and TALE were also predicted to bind to the promoter of *MtCYCC1* (Table 6.5). These four classes have overlapping functions similar to the aforementioned TFs. Depending on similarities in the predicted function of *MtCDK8* and *MtCYCC1* (Table 6.2, 6.3) being part of the CDK8 module as well as similarities of the functions of these predicted TFs, it is highly likely that these classes of TFs are in reality act in regulating functions for the two *M. truncatula* mediator genes. Another key point which was noticed that more than one member for most of the TF classes have been identified for both *MtCDK8* which indicates various members of a particular TF family might act redundantly to regulate function of the aforementioned gene in *M. truncatula*. Pea orthologues of all these TFs have been identified using the pea RNA Seq Gene Atlas database with the assumption of conserved mechanism where the same classes of TFs are likely to regulate the function of *PsCDK8* and *PsCYCC1* genes as well (Table 6.4, 6.5). Further discussion on this issue would be carried out in section 6.5.4.

Table 6.4. Predicted transcription factor binding at promoter of *MtCDK8* (Medtr3g096960) gene generated using Plant Transcriptional Regulatory Map database PlantTFDB 4.0. *MtCDK8* promoter sequence used in this case has been mentioned in appendix sequence A3.1. Pea and *A. thaliana* orthologues were identified through blastP search of *M. truncatula* protein sequences against the Pea RNA seq gene atlas and Phytozyme 12 databases respectively.

Motif	Family	Position	Strand	p-value	Binding sites at *MtCDK8* promoter	Pea orthologue	*A. thaliana* orthologue	Known function	Reference
Medtr7g080460	AP2	30-49	-	6.81E-06	TGGAGCATGGAGGAGAGAGA	PsCam036332	AT5G17430	Regulates flower development, spikelet meristem determinacy, leaf epidermal cell identity, and embryo development	(Nakano et al. 2006)
Medtr1g023660	BBR-BPC	1775-1795	-	2.30E-07	ATCATTTTATCTCTCTCCCTC	PsCam049600	AT5G42520	Regulates ovule identity	(Kooiker et al. 2005)
Medtr7g029400	bZIP	1075-1086	-	5.59E-06	TGCTGACTTGTC	PsCam039836	AT3G30530	Regulates processes such as pathogen defence, light and stress signalling, seed maturation and flower development	(Jakoby et al. 2002)
Medtr7g092750	bZIP	898-912	+	7.72E-06	GATGATGATGTGAGC	PsCam037460	AT5G06839		
Medtr1g053960	C3H	1105-1118	+	2.59E-06	AAAGAAAAAGGGAC	PsCam050891	AT3G12130	Regulates embryogenesis	(Li and Thomas 1998)
Medtr1g056810	Dof	1108-1118	+	6.59E-07	GAAAAAGGGAC	PsCam039066	AT3G52440	Regulates seed storage protein synthesis in developing endosperm, light regulation of genes involved in carbohydrate metabolism	(Lijavetzky et al. 2003)
Medtr4g022370	Dof	571-591	+	2.02E-06	GTAAAAGAGAAAGGAAAAAAA	PsCam037780	AT5G02460		
Medtr8g079060	Dof	1427-1447	-	7.86E-06	ACTCTTTTATTTTTACTTTAG	PsCam037850	AT3G50410		
Medtr2g016030	Dof	1100-1120	+	9.77E-06	TCATCAAAGAAAAAGGGACCA	PsCam040360	AT1G29160		
Medtr5g010940	ERF	756-770	-	8.97E-06	TTATGTCGGCATTCA	PsCam002503	AT1G12610	Regulates hormonal signal transduction, response to biotic and abiotic stresses, regulation of metabolism, and in developmental processes	(Nakano et al. 2006)
Medtr5g010930	ERF	756-770	-	8.97E-06	TTATGTCGGCATTCA	PsCam002549	AT1G12610		
Medtr6g465530	ERF	756-770	+	9.25E-06	TGAATGCCGACATAA	PsCam050192	AT4G25470		
Medtr3g065980	GRAS	569-588	+	7.93E-06	GTGTAAAAGAGAAAGGAAAA	PsCam055871	AT2G01570	Regulates root development, axillary shoot development, and maintenance of the shoot apical meristem	(Bolle 2004)
Medtr8g063600	MYB	382-394	-	6.42E-06	GTATTGGATAAGA	PsCam036568	AT5G08520	Regulates plant secondary metabolism, as well as the identity and fate of plant cells	(Stracke et al. 2001)
Medtr2g064160	MYB	679-699	-	8.84E-06	GTGGCATTTAGGGTGGGTAGG	PsCam011377	AT3G12720		

Table 6.5. Predicted transcription factor binding at promoter of *MtCYCC1* (Medtr7g055650) gene generated using Plant Transcriptional Regulatory Map database PlantTFDB 4.0. *MtCYCC1* promoter sequence used in this case has been mentioned in appendix sequence A3.2. Pea and *A. thaliana* orthologues were identified through blastP search of *M. truncatula* protein sequences against the Pea RNA seq gene atlas and Phytozyme 12 databases respectively.

Motif	Family	Position	Strand	p-value	Binding sites at *MtCYCC1* promoter	Pea Orthologue	*A. thaliana* orthologue	Known function	Reference
Medtr7g080460	AP2	1643-1662	-	6.20E-08	GAGAGAGAAAGAGAGAGAGG	PsCam036332	AT5G17430	Regulates of flower development, spikelet meristem determinacy, leaf epidermal cell identity, and embryo development	(Nakano et al. 2006)
Medtr1g023660	BBR-BPC	1644-1664	+	1.53E-10	CTCTCTCTCTTTCTCTCTCCA	PsCam049600	AT5G42520	Regulates ovule identity	(Nakano et al. 2006)
Medtr7g029400	bZIP	1351-1362	-	9.07E-06	AACTGACGAGGA	PsCam039836	AT3G30530	Regulates processes such as pathogen defence, light and stress signalling, seed maturation and flower development	(Jakoby et al. 2002)
Medtr5g006530	CPP	259-271	+	4.95E-06	CAAAATTTTAAAG	PsCam048966	AT4G14770	Regulates development of reproductive tissue and control of cell division in plants	(Yang et al. 2008)
Medtr2g043050	ERF	581-594	-	5.43E-06	AGGGACGGAGGGAG	PsCam039843	AT1G19210	Regulates hormonal signal transduction, response to biotic and abiotic stresses, regulation of metabolism, and in developmental processes	(Nakano et al. 2006)
Medtr3g065980	GRAS	1639-1658	-	5.72E-08	GAGAAAGAGAGAGAGGGGAA	PsCam055871	AT2G01570	Regulates root development, axillary shoot development, and maintenance of the shoot apical meristem	(Bolle 2004)
Medtr8g089895	HD-ZIP	1768-1778	-	8.13E-06	CAATCATTAAC	PsCam043007	AT2G22430	Regulates organ development, meristem maintenance, responses to environmental conditions.	(Ariel et al. 2007)
Medtr8g033250	MIKC_MADS	1643-1663	+	1.73E-09	CCTCTCTCTCTTTCTCTCTCC	PsCam044484	AT2G45660	Regulates floral organ identity determination	(Pařenicová et al. 2003)
Medtr1g017080	TALE	1644-1663	+	4.02E-08	CTCTCTCTCTTTCTCTCTCC	PsCam043079	AT4G08150	Regulates meristem formation and/or maintenance, organ morphogenesis, different aspects of the reproductive phase	(Hamant and Pautot 2010)

6.5 Discussion

6.5.1 LATE3 and LATE4 are likely to act together in carrying out the same regulatory pathway in a conserved manner

The present study revealed genetic complementarity between *late3-1* and *late4-2* mutants (Figure 6.2) as well as strong physical interaction between LATE3/PsCDK8 and LATE4/PsCYCC1 proteins (Figure 6.3) suggesting that both these genes are likely to act in the same pathway such as floral initiation in a mutually depedent manner. The results are in line with the notion that CDKs such as CDK8 do not have much kinase activity in the absence of associated cyclins such as CYCC1 and active kinase is formed only by CDK-cyclins (Wang et al. 2004; Tank and Thaker 2011).

Previously, CDK8 and CYCC1 were found to interact both physically and functionally with MED12 and MED13 in yeast and animal cells (Borggrefe et al. 2002; Taatjes et al. 2004; Andrau et al. 2006). In the model plant species *A. thaliana*, yeast two hybrid analysis showed that only CYCC1-2 interact physically with CDK8 and MED13 (Ito et al. 2011), however physical interaction between CDK8 and both the CYCC1 proteins was reported in another study by split luciferease complementation assay (Zhu et al. 2014). In the latter study, interaction between CDK8 and CYCC1-2 was stronger than that of CDK8 and CYCC1-1. Besides, the study of Ito et al. 2011 could not establish physical interaction via yeast two hybird assay between MED13 and CDK8/MED12 as well as between CDK8 and MED12. Nevertheless, since CDK8 and CYCC1-2 showed interaction, so it was proposed that CDK8 could interact with MED13 via CYCC1-2. On the other hand, even though MED12 and MED13 did not show physical interaction in this study, functional relationship between these two genes was established in another study carried out by Imura et al., 2012. In this case, *med12* and *med13* mutants showed similar phenotypes such as late flowering and no additive effect of the phenotype of these single mutants for traits such as seedling morphology and flowering time was observed in the *med12 med13* double mutant suggesting that the underlying genes act in the same regulatory pathway (Imura et al. 2012) which was also consistent with the *med12 med13* double mutants in *Drosophila* (Janody et al. 2003; Loncle et al. 2007) as well as *cdk8 cycc1* in *Drosophila* (Xie et al. 2015) and yeast (Banyai et al. 2017). Even though direct physical interaction between MED13 and MED12, MED12 with CDK8/CYCC1 and MED13 with CDK8 has not been shown in plant system yet, however it was found that these components interact physically in

Drosophila which used methodologies such as GST pull down and co-immunoprecipitation assay (Janody et al. 2003; Loncle et al. 2007). Overall, the results of physical and genetic interaction observed for *LATE3/PsCDK8* and *LATE4/PsCYCC1* in the present study is consistent with already known information about such interactions in other systems which indicate a highly conserved role for components of the CDK8 module across eukaryotic system.

The findings of the STRING protein-protein interaction analysis for components of the CDK8 module in *A. thaliana* which was carried out in the current study showed predicted mode of action between MED12 and MED13 only and former gene in turn showed different molecular actions with CDK8, CYCC1-1 and CYCC1-2 (Figure 6.14 A-B), thus MED12 has been inferred to play a central role in mediating the function of entire CDK8 module. Similar central role for MED12 was reported in yeast and human and MED13 was found to be involved in establishing a strong association of CDK8 module with the middle module of core mediator (Tsai et al. 2013; Wang et al. 2013; Jeronimo and Robert 2017) followed by weaker connection established by MED12 (Tsai et al. 2013). Association of core mediator-CDK8 module was found to be eliminated completely by deletion of MED13 (Tsai et al. 2013). A similar trend was found in *A. thaliana* as well since the late flowering phenotype of *med13* and *med12* mutants were more stronger than the *cdk8* mutants (Zhu et al. 2014). The *cycc1* mutants of *A. thaliana* were also found to be late flowering in the current study (section 5.4.7). Besides, the broader network of CDK8 and CYCC1 in *A. thaliana* (at highest confidence interaction score 0.90) included components of the cell cycle and mediator complex (Figure 6.14 C, Table 6.1) which was also in line with known role of these two genes in other systems (Yang et al. 2016; Sage 2004). Since the physical and genetic interaction studies involving *LATE3*/*PsCDK8* and *LATE4*/*PsCYCC1* of pea showed results which are consistent with CDK8 module components in other systems including *A. thaliana*, therefore it can be hypothesized that the function and sturctural organization of the CDK8 module might be conserved in pea as well. Future research in this direction involving already identified components of pea CDK8 module and core mediator (Table 5.4) and orthologous of the *A. thaliana* genes from cell cycle cluster (Figure 6.14 C, Table 6.1) would reveal better understanding in this regard.

6.5.2 LATE3 and LATE4 genes seem to mediate flowering by regulating expression of important pea flowering genes

During the current project, regulatory interaction experiments involved important pea flowering genes which act as florigen and regulators of vegetative and reproductive meristem identity. Among them, *FTb2* is considered the major mobile signal needed for onset of flowering in pea (Hecht et al. 2011; Weller and Ortega 2015), but no obvious difference in expression of this gene was observed in the leaf of *late3* and *late4* mutants compared to WT which potentially rules out that *FTb2* expression suffered any significant impact due to mutation in *LATE3* and *LATE4* genes (Figure 6.6). However, expression pattern of the other mobile signal gene *FTa1* (Figure 6.6) as well as the integrator of leaf bound mobile signal in apex, i.e., *FTc* (Figure 6.6) (Hecht et al. 2011; Weller and Ortega 2015) revealed a positive role of *LATE3* and *LATE4* in regulation of these two *FT* genes. Regulatory interaction studies further showed upregulation of VM identity gene *LF* in the *late3* and *late4* mutants compared to WT and no obvious difference between these three genotypes for I_1M identity gene *DET* (Figure 6.6). Previously, *LF* was found to be epistatic to *FTa1*/*GIGAS* as the *lf gigas* double mutants flowered at the same time with *lf* and thereby suppressed the non-flowering phenotype of *gigas* (Hecht et al. 2011). Therefore, *LATE3* and *LATE4* might suppress *LF* activity via *FTa1* in apex. The very low expression of *DET* in the apex in WT, *late3-1* and *late4-1* genotypes observed in the current study could be due to occurrence of *DET* induction after floral transition as explained earlier (Foucher et al. 2003).

It has been reported previously that floral commencement in pea is closely linked to the attainment of I_1M identity through activity of *DET* in the SAM (Berbel et al. 2012). *DET* represses expression of *VEG1* and *PIM* in the I_1 meristem while *VEG1* (I_2M identity gene) suppress expression of *DET* in I_2 meristem so that *PIM* can be expressed in axillary floral meristems (Berbel et al. 2012). The present study showed that *VEG1* expression reached a peak in WT prior to floral onset at 28 DAS similar to *FTc* (Figure 6.6). In this case, *LATE3* and *LATE4* genes played a positive role indicating potential aberration in both mobile floral signal integration and I_1M/I_2M transition processes in the apex of *late3* and *late4* mutants. In contrast, the expression of *PIM* and *UNI* spiked in WT (Figure 6.6) just prior to development of first visibile flowering bud at 42 DAS suggesting positive role for *LATE3* and *LATE4* in I_2/FM transition. It was shown in a previous studies that *FTa1* promotes expression of *FTc*, *UNI* and

PIM in apex (Hecht et al. 2011). Therefore, it can be speculated that *LATE3* and *LATE4* act upon these genes indirectly via *FTa1*.

The process by which CDK8 module of the mediator complex repress transcription of genes involved in various biological processes have been revealed by detailed structural studies involving yeast and humans and it was found that CDK8 module bind revesibly to core mediator and thereby block transcriptional re-initiation mediated by RNA polymerase II (Elmlund et al. 2006; Knuesel et al. 2009b; Tsai et al. 2013). In yeast, components of the CDK8 module was proposed to compete with the carboxy-terminal domain (CTD) driven RNA polymerase II (RNPII) biding site located in the head and middle module of core mediator. The CTD dependent interaction of RNPII and core mediator is blocked off upon binding of the CDK8 module. In addition to blockage of CTD mediated transcriptional activation, human CDK8 module component was also reported to suppress gene expression by depositing H3K9 and H3K27 epigenetic marks to target genes by working in cooperation with histone methyltransferase G9a (Ding et al. 2008; Chaturvedi et al. 2009; Chaturvedi et al. 2012). Repression of transcription due to disruption in RNA polymerase II-core mediator interaction is considered to be flexible type of transcriptional regulation whereas those imposed by H3K9 and H3K27 are likely cause far more stable suppression. Well known role for H3K9 and H3K27 in seed to seedling and vegetative to reproduction stage transition is already available (Bastow et al. 2004; Bouyer et al. 2011). On the other hand, transcriptional activation role for entire CDK8 module component specifically CDK8 is also known in animals by assisting various processes such as coactivator recruitment, scaffold formation, transcriptional elongation and H3 modification (Larschan and Winston 2005; Donner et al. 2007; Carrera et al. 2008; Galbraith et al. 2010; Nemet et al. 2014).

In *A. thaliana*, mutation in genes belonging to the core mediator of mediator complex such as *MED8*/*18*/*25* resulted in late flowering phenotype under LD and *FT* was downregulated in the respective mutants (Kidd et al. 2009, Zheng et al. 2013). Since the core mediator is known to promote transcription, therefore it is likely that *MED8*/*18*/*25* would positively regulate transcription of *FT* to regulate flowering. On the other hand, mutation in CDK8 module genes of the mediator complex such as *MED12/MED13* also caused late flowering phenotype under LD and *MED12* was reported to promote expression of *FT*, *FUL*, *AP1* and *LFY* orthologues of which have been studied in the current study (Imura et al. 2012). In another study, *MED13*

was reported to promote expression of *FT* (Gillmor et al. 2010). These results and that of the present study involving *PsCDK8*/*PsCYCC1* strongly guides towards the idea of conservation of function of various CDK8 module genes across plant species with specific relevance to flowering. Therefore, considering a conserved molecular mechanism of function of CDK8 module in eukaryotes, i.e., repression of transcription, *LATE3*/*PsCDK8* and *LATE4*/*PsCYCC1* may promote flowering by blocking transcription of some unknown intermediary gene (or genes) which in turn repress the function of *FTa1*, *FTc*, *UNI*, *VEG1*, *VEG2* and *PIM*. The repression of the intermediaty gene (s) could occur either by preventing re-initiation carried out by relevant pea RNA polymerase II or by accumulating aforementioned/other epigenetic marks. Similarly, *LATE3* and *LATE4* may negatively regulate transcription of *LF* either directly or by promoting activity of intermediary gene (s) that suppress *LF* function. In contrast, *LATE3* and *LATE4* being complementary genes could also act together in transcriptional activation of *FTa1*, *FTc*, *UNI*, *VEG1*, *VEG2* and *PIM*. Under such situation, *LATE3*/*PsCDK8* and *LATE4*/*PsCYCC1* may attenuate *LF* transcripton by stimulating transcription of some intermediary gene (s) that repress *LF* function. Moreover, *LATE3* and *LATE4* may suppress activity of the potential intermediary target gene (s) epigenetically and thus promote expression of aforementioned positive regulators of flowering whereas such epigenetic repression could be exerted directly upon negative regulator of flowering such as *LF*. Transcription regulation by mediator complex and histone modification by methyltransferases are known to be involved in universal regulation of various developmental processess in plants (Kidd et al. 2011; Thorstensen et al. 2011). So, both of these complexes might act together in regulation of flowering in pea with specific relevance to *LATE3*/*PsCDK8* and *LATE4*/*PsCYCC1*. In the previous chapter, potential components of both these major complexes in pea have been identified (Table 5.3, 5.4). Therefore, future studies involving pea CDK8 module and core mediator components as well as histone modifiers is likely to unveil further understanding in this direction.

The present study also revealed that stronger alleles of *late3* and *late4* mutats flower in a photoperiod insensitive manner (Figure 6.4). Previously, it was found that *med12* and *med13* mutants in *A. thaliana* show photoperiod response and both *MED12* and *MED13* was found to promote expression of key photoperiodic signal integrator *CO* (Imura et al. 2012; Gillmor et al. 2010). In pea, any key photoperiodic signal integrator in the leaf similar to *CO* has not

been identified yet (Weller and Ortega 2015). However, *A. thaliana* orthologues of the red, far-red and blue-light photoreceptors namely phyB, phyA and cry1 involved in photoperiod regulated flowering have been identified in pea (Weller and Reid 1993; Weller et al. 1997a; Weller et al. 2001; Platten et al. 2005). Few other orthologues of *A. thaliana* photoperiod and circadian clock mediated flowering genes are *LATE1*/*PsGI, DNE*/*PsELF4* and *LATE2*/*PsCDFc1* where the respective mutants were photoperiod insensitive (Hecht et al. 2007; Liew et al. 2009; Weller and Ortega 2015; Ridge et al. 2016). As discussed above, regulatory interaction studies revealed that *LATE3*/*PsCDK8* and *LATE4*/*PsCYCC1* do not regulate expression of *FTb2* whereas they positively regulate expression of *FTa1* in leaf (Figure 6.4 B) which is likely to act downstream of the yet to be identified key pea photoperiod integrator gene, however it is not yet clear whether such a process is actually present in pea (Weller and Ortega 2015). Expression of *FTa1* in leaf was found to be mediated by photoperiod previously and it was proposed that *LATE1*/*PsGI* controls *FTa1* expression via *FTb2* under LD whereas *DNE*/*PsELF4* regulates expression of *FTa1* independently of *FTb2* under SD (Hecht et al. 2011). Moreover, it was revealed that consistent with *FT* regulation in *A. thaliana* by FKF1/GI/CDF module, a similar pea module comprised of PsFKF1/LATE1(PsGI)/LATE2(PsCDF1c) regulates *FTa1* expression in the photoperiod pathway independently of *CO* like genes (Ridge et al. 2016). As *LATE3*/*PsCDK8* and *LATE4*/*PsCYCC1* belongs to global transcriptional regulator mediator complex, therefore it can be speculated that they may regulate expression of *FTa1* via the known pea photoperiodic flowering regulators, i.e., *PHYA*, *PHYB*, *DNE*, *LATE1, LATE2, FKF1* and future interaction studies could test this hypothesis.

In pea, the process of vernalization mediated flowering is not understood and only little information is available until now where the *FTa1* gene is probably important for response towards vernalization (Weller and Ortega 2015; Beveridge and Murfet 1996). In the present study, it was observed that *late3* and *late4* mutants exert no response to vernalization under both LD and SD (Figure 6.5). The results of vernalization response in WT NGB5839 (generated from cv. Torsdag) under both LD and SD was consistent with that of previously reported for cv. Virtus (Beveridge and Murfet 1996). In *A. thaliana*, it was reported that *AtMED12* and *AtMED13* suppress expression of the key vernalization mediated flowering pathway gene *FLC,* but no experiment on the vernalization response of *med12* and *med13* mutants was conducted in these studies (Gillmor et al. 2010; Imura et al. 2012). Since the role of *AtCDK8*

and *AtCYCC1* is not yet known in the vernalization process, therefore future research in *A. thaliana* could reveal whether they also regulate *FLC* expression or not. Besides, checking vernalization response for the mutants of all the genes of AtCDK8 module would assist in linking whether differential expression of *FLC* as observed for *med12* and *med13* is related to altered response towards vernalization treatment. Moreover, it would also be of interest to check the expression of *FTa1* gene in *late3* and *late4* mutants upon vernalization treatment in pea.

6.5.3 Investigation of genetic interaction with LF, DET and SN

During the present study, genetic interaction of *lf* with *late3* and *late4* revealed that *LATE3* and *LATE4* are hypostatic to *LF* since the *lf late3-2* and *lf late4-1* flowered at around the same time like *lf* single mutants (Figure 6.9 A-C). These results were consistent with that of regulatory interaction as mutation in *LATE3* and *LATE4* genes caused upregulation of the *LF* in the *late3* and *late4* mutants compared to WT. Thus, these combined findings from two separate experiements support the upstream regulatory activity of *LATE3* and *LATE4* in flowering initiation through repression of the expression of *LF*. The genetic interaction studies of *late3-2* and *late4-1* with *det* had to be carried out using *lf det late3-2* and *lf det late4-1* triple mutants as no *det late3-2* and *det late4-1* mutants were available for this purpose. The *lf det* mutant is known to exhibit additive effect with no extra phenotype which means that *LF* and *DET* do not act redundantly, thus together they represent *A. thaliana TFL1* function (Foucher et al. 2003; Murfet 1989). In the current project, the *lf det late3-2* and *lf det late4-1* triple mutants showed more acute form of determinate growth possessing malformed flowers along main stem compared to *det* and *lf det* mutants where the shoot apices ended in a normal terminal flower eventually transforming into pods (Figure 6.8 A, 6.10 A-C). These results strongly suggest that there is additive effect of *late3* and *late4* mutations when added to *det* in relevance to apical meristem determinacy. However, *LATE3* and *LATE4* act upstream of *DET* and *LF* for determining floral initiation and length of reproductive phase since the *lf det* flowered eariler and had shorter reproductive phase compared to *lf det late3-2* and *lf det late4-1* triple mutants.

On the other hand, investigation on genetic interaction between circadian clock mutant *sn4* and *late4-1* revealed that *SN*/*PsLUX* is hypostatic to *LATE4* to mediate flowering since the *sn late4-1* double mutants flowered at around the same time as *late4-1* (Figure 6.13 A-B). *SN* is

known to be involved in photoperiod response and was found to repress expression of key pea flowering genes such as *FTa1*, *FTc* and *PIM* (Liew et al. 2014). In the current study, *LATE3* and *LATE4* were also found to be involved in photoperiod response. Hence, results from both the genetic interaction and photoperiod experiments suggest that *SN* is likely to act upon *FTa1*, *FTc* and *PIM* through *LATE4*. Unfortunately, it was not possible to recover *sn late3-1* double mutants during the current project due to lack of segregation. However, as the phenotypes of *late3* and *late4* was more or less similar and the underlying genes showed genetic complementarity and strong physical interaction to carry out the same regulatory process, therefore it can be hypothesized that role of *LATE3* is likely to have an analogous genetic interaction with *SN*.

Apart from *SN*, three other evening complex genes namely *HR*, *PPD* and *DNE* have been identified which are pea homologues of *PsELF3a*, *PsELF3b* and *PsELF4* respectively (Liew et al. 2009; Weller et al. 2012b; Liew et al. 2014; Rubenach et al. 2017). Genetic interaction studies revealed miminal effect on flowering in *ppd-3* and *dne-1* mutants in the *HR* background suggesting redundant role of *HR*/*PPD*/*DNE* genes in regulating floral initiation (Liew et al. 2014; Rubenach et al. 2017). In contrast, *sn4* showed epistatic relationship with *hr* for onset of flowering and expression of *SN* was found to be regulated by *HR* (Liew et al. 2014; Rubenach et al. 2017). Therefore, it can be speculated that these three genes are likely to act upon *LATE3* and *LATE4* in a similar way to *SN*. In another genetic interaction studies, the *LATE1*/*PsGI* showed hypostatic relationship with *SN* for the regulation of flower initiation (Hecht et al. 2007). Therefore, *LATE1* might also act upon *LATE3*/*LATE4* in a similar way like *SN* for controlling onset of flowering. Future genetic interaction studies would make the scenario more clearer.

6.5.4 Function and regulation of CDK8 and CYCC1 is predicted to be conserved between M. truncatula and pea

The AraNET database was used in the current study for predicting functions of *CDK8* and *CYCC1* gene in *M. truncatula* as it has published genome sequences and was used as a tool of comparative genomics based experimentations in pea. The inferred functions obtained from this analysis based on mutant phenotype, genetic and physical interaction data of network neighbours of *MtCDK8* and *MtCYCC1* revealed potential role of both these genes in flowering, different types of vegetative and reproductive organ formation, transition from vegetative to

reproductive stage and response to abiotic and biotic stress (Table 6.2, 6.3). These results are consistent with what was observed based on proper scientific experimentations for *LATE3 PsCDK8* and *LATE4/PsCYCC1* in pea and *AtCYCC1* in *A. thaliana* during the current study. Besides, previous characterization of *cdk8*, *med12* and *med13* mutants in *A. thaliana* showed similar functions of the underlying genes (Gillmor et al. 2010; Imura et al. 2012; Wang and Chen 2004; Zhu et al. 2014). Likewise, research using components of the core mediator such as *AtMED25*, *AtMED8*, *AtMED18* also provided evidence for such functions (Kidd et al. 2011; Samanta and Thakur 2015). Role in flowering and other developmental traits for *AtMED15*, *AtMED16*, *AtMED17* and *AtMED20* have also been reported (Malik et al. 2017). In addition, involvement of *CDK8*, *MED12*, *MED13* and *CYCC1* in relevance to biotic stress tolerance in *A. thaliana* was established previously (Zhu et al. 2014). Moreover, *CDK8* and other core mediator complex components was found to mediate various abiotic stress tolerance in *A. thaliana* (Kidd et al. 2011; Ng et al. 2013; Samanta and Thakur 2015). Furthermore, expression profile of *MtCDK8* and *MtCYCC1* genes across various vegetative and reproductive tissues was found to be similar to that of pea (chapter 5, appendix figure A2.3 D-E, A2.4 D-E). Altogether, these predicted results on functions of *MtCDK8* and *MtCYCC1* provide hints that the functions of mediator complex genes are probably conserved between various plant species. Results obtained from AraNET v.2 could be validated in future by characterizing *Mtcdk8* and *Mtcycc1* mutants after ordering relevant seeds from *M. truncatula* tnt mutant database (https://medicago-mutant.noble.org/mutant/).

During the present study, usage of PlantTFDB database led to projecting that *CDK8* and *CYCC1* genes in *M. truncatula* and pea are likely to be regulated by TFs belonging to AP2, BBR-BPC, DOF, ERF, bZIP, MYB etc. groups (Table 6.4, 6.5). In general, these TFs mostly have overlapping functions related to plant development and biotic and abiotic stress response which are in line with the function of *PsCDK8* and *PsCYCC1* genes revealed in the current study. In a previous study, analysis of the promoter of all the mediator genes in *A. thaliana* showed presence of different responsive elements important for hormonal and various abiotic stresses (Pasrija and Thakur 2012) which indicates that relevant (yet unknown) transcription factors are likely to act upon these genes in order to regulate their function for the aforementioned regulatory processes. So, the identified TFs from the present study may in reality control functions of *CDK8* and *CYCC1* in *M. truncatula* and pea. The binding motifs of

the TFs that have been predicted to regulate expression of *CDK8* and *CYCC1* gene in *M. truncatula* and pea could be validated in future by Chip-PCR. Besides, expression of *CDK8* and *CYCC1* could be checked in over-expressed and knock out line for respective TF genes to reveal potential differential regulation of those genes. A combinantion of these three approaches were applied for validating three computationally predicted TFs regulating function of *Sestrin3 (SESN3)* gene in human which play role in glucose and lipid metabolism as well as defense against oxidative stress (Srivastava et al. 2016). Moreover, the functional significance of regulation could be revealed through transient co-transfection studies by using various combinations of *CDK8* and *CYCC1* promoter deletion mutant-reporter vector along with respective TF expression vector which is likely to determine the effect of TF on the transcriptional activity of *CDK8*/*CYCC1* as well as the specific regulatory region of promoter needed for binding of the TF. Similary, co-transfection using full length *CDK8*/*CYCC1* promoter-reporter construct along with specific TF deletion mutant-overexpression vector (e.g. deleted *CDK8*/*CYCC1* binding domain of the TF) would potentially reveal the effect of the specific binding motif of TF on the transcriptional activity of the *CDK8*/*CYCC1* genes. Successful execution of these two methods along with aforementioned Chip-PCR and overexpression/knock out line generation revealed that transcriptional activation of *OCT4*, a human gene involved in induction and maintenance of cellular pluripotency is mediated by an ETS transcription factor namely *PEA3* (Park et al. 2014).

6.5.6 Concluding remarks

Results of the current chapter showed that LATE3/PsCDK8 and LATE4/PsCYCC1 proteins interact physically and they act in the same regulatory pathway in a mutually dependent manner which was consistent with previous similar findings on different components of CDK8 module of the mediator complex in other systems including *A. thaliana*. The function of the CDK8 module is likely to be conserved in plant system as various members of cell cycle and mediator complex have been identified as interacting partners of AtCDK8, AtCYCC1-1 and AtCYCC1-2 through STRING database. Besides, *LATE3*/*PsCDK8* and *LATE4*/*PsCYCC1* seemed to promote photoperiod mediated flowering most likely by controlling expression of mobile signal regulating gene *FTa1* in leaf and they also assist in florigen integration in apex by regulating *FTc* expression. Likewise, they regulate acquisition of inflorescence and floral

meristem identity by promoting expression of *VEG1*, *VEG2*, *UNI* and *PIM* and repressing expression of *LF*. In addition, pea evening complex gene *SN*/*PsLUX* is likely to repress flowering by acting negatively upon *LATE3* and *LATE4* as suggested by genetic interaction studies between *sn4* and *late4-1*. Moreover, a systems biology based approach was undertaken via usage of AraNET database in order to predict function of *MtCDK8* and *MtCYCC1* genes which showed high similarities with that of *PsCDK8*, *PsCYCC1* and *AtCYCC1* revealed in the present study so far, thereby strengthening perception that this particular database can be used in greater confidence in order to generate hypothesis about the function of unknown genes. Furthermore, exploitation of PlantTFDB database unveiled predicted TFs regulating functions of *MtCDK8* and *MtCYCC1* which suggest that both these genes are controlled by TFs belonging same groups, e.g. AP2, ERF, bZIP which are known to have functions similar to that of *MtCDK8* and *MtCYCC1*. This work laid down the foundation for future investigations in order to determine the crucial cis-elements that regulate function of these two universal regulatory genes in pea and *M. truncatula*.

Chapter 7: Regulation of response to various environmental factors by *LATE3* and *LATE4* loci

7.1 Introduction

During the entire life cycle starting from germination of seeds to senescence, plants are highly affected by various environmental factors. As plants are sessile and are unable to run away like animals to more favorable conditions upon subjection to adversity, so they have evolutionarily developed strategies to survive under such atmosphere in order to complete vegetative and reproductive stages of their lives successfully. In this case, different types of interconnected molecular response at the level of nucleic acids, proteins and metabolites are triggered within the plant cells to perceive and then initiate response towards a particular environmental factor which ultimately determine the phenotypic outcome (Soni et al. 2015). Role of mediator complex in relation to regulating response to various developmental, biotic and abiotic factors in plants is already known (Kidd et al. 2011; Samanta and Thakur 2015; Yang et al. 2016; Dwivedi et al. 2017; Malik et al. 2017). *AtMED25* have been reported to control salinity (Elfving et al. 2011), drought stress response while *AtMED2*, *AtMED14* and *AtMED16* was found to be important for cold acclimation (Hemsley et al. 2014). Likewise, role of *AtCDK8* for generating response to H_2O_2 and cold stress was unveiled and this gene was proposed as a switch between growth and stress response (Ng et al. 2013). In addition, components of the CDK8 module in *A. thaliana* was found to be vital for regulating response towards biotic stress (Zhu et al. 2014). Moreover, expression of all the components of the mediator complex in *A. thaliana* seedlings is downregulated under continuous darkness and upregulated except for *MED15* and *MED37* under high intensity white light (800 μ E m^{-2} s^{-1} for 3 minutes) compared to normally grown seedlings (14 hour light, 10 hour dark cycle) (Pasrija and Thakur 2012). Similarly, salt stress causes upregulation of all of the *A. thaliana* mediator genes (Pasrija and Thakur 2012). Since various mediator complex components were found to regulate response to different abiotic factors in *A. thaliana* and *LATE3* and *LATE4* were identified as *PsCDK8* and *PsCYCC1* of the pea mediator complex and system biology study on *MtCDK8* and *MtCYCC1* (Table 6.2 and 6.3) showed potential function of these genes in this

direction, therefore checking the response of *late3* and *late4* mutants towards different environmental factors likely to reveal whether the underlying genes are imporatant for growth and adaption in pea.

Light is one of the most essential exogenous factor that governs the pattern of growth and development of photoautotrophs like plants (Han et al. 2007; Josse and Halliday 2008; Arsovski et al. 2012). Such type of growth through activation of light meditated signalling cascades is defined as photomorphogenesis. Post seed germination seedling development phase is primarily determined by photomorphogenesis and proper course of action by the seedling is needed during this time which would ultimately decide photosynthetic as well as reproductive success (Mohr 2012). In the complete absence or limited availability of light, plants exhibit growth patterns suitable to dark conditions which is termed as skotomorphogenesis. The principles behind dark phenotypes is to make proper use of the limited nutrients available in the seed which would assist the plant to reach towards light. Such phenotypes include different etiolated characteristics such as pronounced apical hook formation in dicots or coleoptile in monocots, prompt elongation of stem, absence of leaf as well as cholorophyll production and minimum root growth (Arsovski et al. 2012; Josse and Halliday 2008). On the other hand, photomorphogenic growth is exhibited by some trademark de-etiolated traits such as unlatched apical hook or coleoptile, suppression of stem elongation, expanded leaf as well as cholorophyll generation and proper root development (Nemhauser and Chory 2002; Arsovski et al. 2012). Plants have evolved various photoreceptors for initiation and subsequent regulation of growth under the available light condition. In the model plant species *A. thaliana*, there are five differenet photoreceptors namely Phytochrome A-E (Whitelam and Devlin 1997) which are used by the plant to perceive red/far-red light (600 - 750 nm). Likewise, the model plant species possess seven different photoreceptors to recognize blue light (450 - 495 nm) which are as follows: two cryptochromes (CRY1 and CRY2), two phototropins (PHOT1 and PHOT2) and three Zeitlupe family members (ZTL, FKF1 and LKP2) (Kami et al. 2010; Tilbrook et al. 2013).

UV radiation is a factor that is responsible for evolutionary origin of different species as it causes mutation in DNA. It also affects the growth and development of plants. Among the various components of sun light, UVB lies within the wavelength of 280-315 nm. Almost all of the harmful UVC (<280 nm) as well as much of the UVB radiation is largely hindered from

reaching the surface of the earth by the protectant stratospheric ozone layer (Heijde and Ulm 2012). Solar UV radiation that still reaches earth surface consists mostly of less harmful UVA (315 - 400 nm) and only part of UVB. Overall, UVB contributes to less than 0.5% of the total light energy that reaches the earth surface. However, plant undergoes various morphological changes due to exposure to UVB such as generation of thicker leaves, smaller petioles, reduction in stem elongation, modulation in leaf and inflorescence architecture as well as root : shoot ratio and enhanced tillering (Jansen 2002; Robson et al. 2003; Furness et al. 2005; Hectors et al. 2007; Yang et al. 2008; Wargent et al. 2009a; Wargent et al. 2009b; Klem et al. 2012). In *A. thaliana*, UVB defence mechanism is initiated by recognition of these raditions by a β-propeller protein photoreceptor named as UVR8 (Kliebenstein et al. 2002; Rizzini et al. 2011).

Temperature plays a significant role in growth and development of plants . During an era of rapid climate chage, rise in global temperature is a matter of high concern for scientists which requires constant innovation of proper mitigation strategies. Certain temperature regime falling out of the optimum range on either side could exert stress upon the plant and this range could vary from one species to another. Plants are subjected to cold stress upon exposure to 0-15° C (Yadav 2010; Sadras et al. 2012; Bueckert et al. 2015; Abesingha 2015). Heat stress is another temperature related environmental factor that could render significant negative impact on crop production. The optimum temperature for growth of pea is 15-20° C (Mahony 1991; Stanfield et al. 1966) and this growth is significantly reduced at 30° C and above (Pate 1977; McDonald and Paulsen 1997). In pea variety Alaska, rate of progress towards flowering increased persistently from 6° C to 24° C (Summerfield and Roberts 1988). In another pea variety cv. Carneval, temperature over 25° C reduced yield (Sadras et al. 2012; Bueckert et al. 2015) and heat stress of 33-35° C (6 hours each day for 4 days) negatively influenced seed development at maturity (Abesingha 2015). Exposure to 40° C for 2.5 minutes caused serious damage to the photosynthetic system, e.g. thylakoid activity in pea (McDonald and Paulsen 1997). Heat stress could affect the plants adversely in many ways such as generation of excessive reactive oxygen species (ROS), inhibition of seed germination, reduced plant growth, modulation in photosynthesis, phenology and dry matter partitioning, loss of water, reduction in yield and crop quality (Hasanuzzaman et al. 2013). Plant's response to heat stress involves change in expression of genes which encodes osmoprotectants,

detoxifying enzymes, transporters and regulatory proteins (Semenov and Halford 2009; Krasensky and Jonak 2012). HEAT SHOCK TRANSCRIPTION FACTOR A1s (HsfA1s) is known as the master regulator of the transcriptional regulatory network that mediate heat stress response in *A. thaliana* and it is considered as an indicator for heat stress related investigations (Ohama et al. 2017). HsAF1s is hypothesized to regulate expression of various heat stress responsive TFs such as DEHYDRATION RESPONSIVE ELEMENT BINDING PROTEIN 2A (DREB2A), HsfA2, HsfA7a, HsfB and MULTIPROTEIN BRIDGING FACTOR 1C (MBF1c) (Yoshida et al. 2011). Heat shock proteins (HSP) such as HSP70 and HSP90 repress HsfA1 activity under non-stress conditions whereas HsfA1 is released from the suppression upon exposure to heat stress (Yamada et al. 2007; Hahn et al. 2011).

Ambient temperature can be defined as the range which is not high enough to casue heat stress for a certain plant species, yet could initiate modulation in a range of morphological changes known collectively as thermomorphogenesis (McClung et al. 2016). Some of the traits at 27° C in *A. thaliana* resemble to those of the shade avoidance symptoms such as elongated hypocotyls and petioles. Besides, slight increase from 23° C to 27° C is good enough induce flowering in *A. thaliana* under non-inductive short day conditions (Balasubramanian et al. 2006). Fluctuation in ambient temperatures also lead to alteration in various developmental and physiological traits in plants such as proportion of net CO_2 consumption and transpiration, chlorophyll and proline content, number of chloroplasts, size of mitochondria, leaf structure, seed yield etc. (Smillie et al. 1978; Ozturk and Szaniawski 1981; Cornic and Ghashghaie 1991; Todorov et al. 2003; Prasad et al. 2006; Jin et al. 2011). Thus, ambient temperature could impose striking detrimental effect on various traits which are important for crop productivity. Basic helix-loop transcription factors namely PHYTOCHROME INTERACTING FACTOR4 (PIF4) and PIF5 which play role in photomorphogenesis pathway are also known to be crucial components of thermomorphogenesis (Arsovski et al. 2012; Wigge 2013; Quint et al. 2016). Discussion on role of ambient temperature on flowering in plants has also been made in chapter 1.

Salinity is currently a major challenge of agriculturalists that is threatening crop productivity specifically in arid and sem-arid climatic areas (Acosta-Motos et al. 2017). Soil salinity is defined as a matter of concern when it crosses the electric conductivity (EC) threshold of 4 dS/m, i.e., equivalent to 40 mM NaCl that could reduce yield of most crops. Thus, an estimated

800 million hectares of arable lands worldwide are affected by the raising salinity of the soil (Munns and Tester 2008). Toxicity caused by accumulation of NaCl causes increase in soil resistance, reduction in root growth as well as water movement via root, decrease in carbon assimilation rate, leaf abscision etc. (Acosta-Motos et al. 2017). Various researches to date have identified four different mechanisms that are being exploited by salt tolerant euhalophytes under toxic saline conditions which are as follows i. salt exclusion: blocking salt entrance into vascular system, ii. salt elimination: different salt secreting glands and hairs stongly eliminate salts keeping its concentration below a specific threshold, iii. salt succulence: concentration of the salt could be kept constant if the cell storage volume rises constantly via uptake of salt, iv. salt redistribution: high salt concentration generated in the leaf could be diluted by transportation of Na^+ and Cl^- to another location within the plant body. Various genes contribute towards mitigation of plants once toxicity is generated by a certain salt concentration. Among these, tonoplast localized Na^+/H^+ exchanger 1 (NHX1) performs crucial role in salinity detoxification by compartmentalizing toxic Na^+ into the vacuoles in *A. thaliana* (Blumwald and Poole 1985; Deinlein et al. 2014). Likewise, a Na^+/H^+ antiporter named as SALT OVERLY SENSITIVE 1 (SOS1) acts by exporting toxic Na^+ out of the cell (Qiu et al. 2002; Yamaguchi et al. 2013). Another major player in this regard is class I high affinity H^+ transporter (HKT1), a Na^+- K^+ co transporter which carries out removal of Na^+ from the xylem and increases uptake of K^+, thereby high K^+/ Na^+ ratio can counteract toxicity caused by Na^+ (Sunarpi et al. 2005).

Mechanical injury or wound can hamper plants and ultimately yield outcome by providing an window for significant loss of nutrients and entrance of pathogenic microbes leading to disease outbreak in a field condition. Under such situations, plants exploit the same mechanism by which they deal with invading pathogens (León et al. 2001; Vasyukova et al. 2011; Savatin et al. 2014). This involves activation of both local and systemic innate immunity system within the plant through recognition of endogenous damage associated molecular patterns or DAMPs. A peptide named as systemin act as a major wound associated DAMP (Jacinto et al. 1997; Narvaez-Vasquez and Ryan 2004; Schilmiller and Howe 2005; Savatin et al. 2014). In the leaf of *A. thaliana*, an alert mechanism is induced soon after perceiving the danger posed by DAMPs that leads to modulation in wound-activated surface potential as well as generation of reactive oxygen species (ROS) waves which causes priming of the non-

infected parts of the plant body (Mittler et al. 2011; Suzuki and Mittler 2012; Savatin et al. 2014). Besides, long distance defence response involve activation of genes of the jasmonic acid (JA), abscisic acid (ABA) and ethylene biosynthesis pathway (Penacortes et al. 1995; Bergey et al. 1996; Bouquin et al. 1997). Moreover, deposition of callose, suberin, lignins or various phenolics around the site of wound at a later stage act as potential barrier against any potential damages that could occur via microbial infection (Savatin et al. 2014).

Plants living under a certain habitat such as arid area often need to deal with multiple environmental stress simultaneously (Sewelam et al. 2014). Under salt, osmotic and heat stresses, plants respond the activating unique signalling cascade that is different than the process stimulated by individual stresses. In *A. thaliana*, some key genes that are differentially expressed under aforementioned stresses include those from heat shock proteins (HSPs), heat shock regulators and late embryogenesis abundant proteins (LEAs) (Sewelam et al. 2014). Besides, analysis of publicly available AtGenExpress database (Kilian et al. 2007) for wounding, genotoxic, oxidative, UVB, osmotic and salt stress response identified genes of the plant core enviromental stress response (*PCESR*) (Hahn et al. 2013). Among these, a putative zinc finger transcription factor *ZAT12* is regulated by light, oxidative, heat, UVB, ROS, cold and drought stresses and therefore, status of this gene is considered as a suitable marker for research involving those abiotic stresses in plants (Kilian et al. 2007; Davletova et al. 2005; Sakamoto et al. 2004; Hahn et al. 2013). Other examples of such genes is *ZAT10*, heavy-metal-transport superfamily gene *ERD9* and methyl-esterase *MES9*. Both *ZAT12* and *MES9* are known to be repressed by the deposition of H3K27me3 trimethylation mark in the shoot tissue and upregulation of these genes after exposure of stress is probably linked to chromatin remodelling at the aforementioned important loci (Zhang et al. 2007; Zhang et al. 2006; Hahn et al. 2013). MYB-like TFs were found to be upregulated upon simultaneous exposure to drought and heat stresses which suggested their role in regulating the transcription of genes towards both abiotic factors (Rizhsky et al. 2004). The *AtMED25* has been reported to interact with various TFs such as ZFHD1, DREB2A and MYB like proteins which are activated by salt, drought and heat stresses repectively indicating that *AtMED25* is likely a hub for responding to multiple enviromental signals (Elfving et al. 2011; Iñigo et al. 2012; Balderas-Hernández et al. 2013).

7.2 Chapter aim

The overall goal of the current chapter was to understand whether *LATE3* and *LATE4* genes are involved in mediating phenotypic response towards various environmental factors in pea. As part of that, responsiveness of *late3* and *late4* mutants towards development related factors such as different light conditions and ambient temperature was monitored via collection of phenotypic data. Similarly, both the mutants were exposed to abiotic stress factors, e.g., UVB, heat, mechanical injury and salt and information on various traits were compiled in order to reveal their response towards such environmental constraints.

7.3 Materials and methods

Various materials and methods used specifically for this chapter are mentioned here. The general materials and methods exploited in this chapter are given in chapter 2.

7.3.1 Analysis of AtCDK8 and AtCYCC1 gene atlas data

The gene expression profile of *AtCDK8*, *AtCYCC1-1* and *AtCYCC1-2* upon different abiotic stress treatment have been compiled from the AtGenExpress Visualization Tool (AVT) (http://jsp.weigelworld.org/expviz/expviz.jsp) by choosing "AtGE Abiotic Stress" as experiment, "mean-normalized" as the type of normalization and using *AtCYCC1-1* (AT5G48640), *AtCYCC1-2* (AT5G48630) and *AtCDK8* (AT5G63610) accession numbers.

7.3.2 Photomorphogenesis experiment

Plants were grown in separate growth chambers under continuous darkness, white, red, far-red and blue lights at the School of Natural Sciences, University of Tasmania at 20° C. The source of white light was cool-white fluorescent tubes (L40 W/20 S cool white, Osram Germany) at an irradiance of 120 to 140 µmol m^{-2} s^{-1}. On the other hand, Heliospectra RX30 LED lighting system was used having blue (450nm), red (660 nm) and far-red (735 nm) channels at an irradiance of 10 µmol m^{-2} s^{-1}.

After 14 days of growth, leaf area was estimated for each genotype as the product of width and length of a single leaflet at node three. Besides, length of internode was calculated between node one and three from equal numbers of plants.

7.3.3 Ambient temperature experiment

For this experiment, plants were grown at 15° C, 20° C and 25° C. temperatures under LD (16 hour photoperiod) condition in the controlled growth chambers of School of Natural Sciences, University of Tasmania. Information on node of flowering initiation (NFI) were collected from each group of plants if they flowered.

7.3.4 Abiotic stress experiment

In this case, an initial pilot experiment was carried out using only the WT genotype which was grown for 14 days at 20° C, LD condition (16 hour photoperiod) in the controlled growth chambers at the School of Natural Sciences, University of Tasmania. On day 14, various treatments were applied for each of the stress factors and after that plants were allowed to

grow for 14 more days. On day 29th, stem and leaf fresh weight were measured. Likewise, stem length was measured as the length from node one to the bottom of shoot apex. In order to measure the dry weight of the leaves and stem, the respective tissue samples were dried in an oven for 96 hours at 70° C.

Based on the findings of the pilot experiment, a second round experimentations were carried out using WT and one mutant allele from *late3* and *late4* mutants because of seed as well as space limitations in the growth chambers. These plants were also grown for 14 days at 20° C (16 hour photoperiod) in the controlled growth chambers for all stresses apart from salt at the School of Natural Sciences, University of Tasmania. On day 14, various treatments were applied for UVB, heat and wound stress and plants were grown for two more weeks in normal 20° C chamber. After that, phenotypic data were collected for internode length, leaf area, petiole length and shoot dry weight. Internode length was measured as the length between node one and node eight. Leaf area was calculated as the product of length and width of a single leaflet at node eight while petiole length means the length of petiole of the leaf at the same node. Shoot dry weight was measured in the same manner mentioned above. For salt stress, application of treatment and phenotypic data collection process was different.

Various treatments that were used during pilot and second round of experimentations are provided below:

7.3.4.1 UVB stress

Pilot experiment – Plants were subjected to UVB radiation (wavelength 280-315 nm) for 15 and 30 minutes respectively in an UVB chamber (Omron, Japan). Control plants were kept in the growth chamber at the same time.

Second experiment - Respective group of plants from each genotypes were subjected to UVB radiation for 40 and 80 minutes and control plants did not receive any treatment.

7.3.4.2 Heat stress

Pilot experiment- Plants were exposed to 38° C for 1, 2, 3 and 4 hours. Control plants were kept in the normal growth chamber.

Second experiment – Plants were given heat treatment by subjecting them to 45° C for 5 and 8 hours. Control plants remained in the normal growth chamber.

7.3.4.3 Wound stress

Pilot experiment: Two different treatments were applied using a thumbtack: treatment 1 - 10holes on leaf three (five holes/leaflet) and treatment 2 - 10 holes on leaf two and three (five holes/leaflet). Control plants did not receive any treatment.

Second experiment – Two treatments were given using a thumbtack, treatment 1 - 40 holes on leaf two and three (20 holes/leaflet); treatment 2 - 80 holes on leaf two and three (40 holes/leaflet). No treatment was given to the control plants.

7.3.4.4 Salt stress

Pilot experiment: Three different salt treatments consisting of 50 mM, 75 mM and 100 mM NaCl in 500 ml total volume of 1% nutrient solution were applied to the treatment plants. Control plants were applied 500 ml of 1% nutrient solution only.

Second experiment – The second round of experiment was designed for 12 days so that phenotypic data on roots can be collected before the primary root hits the bottom of the 140 mm long pots. From day 6 to 11, the respective genotypes were given two different salt treatments which consisted of 250 mM and 500 mM NaCl in total volume of 100 ml water. Untreated plants were given only 100 ml of water during this time. On day 13^{th}, primary root length was measured as the length from cotyledon to the tip of primary root whereas internode length was measured as the length between node one and node two. In order to measure the dry weight of shoot and root, the respective tissue samples were dried in an oven for 96 hours at 70° C.

7.3.5 Statistical analysis

For photomorphogenesis, UVB, heat, wound and salt experiments, box-whisker plots were generated using R software version 3.3.2. The Upper whisker represents highest 25% of the values whereas lower whisker indicates lower 25% of the values. The box represents middle 50% of the values falling within interquartile range (Q1 to Q3) and the line within the box shows median value. One-way ANOVA followed by Tukey test was performed to determine the level of significance between control and treated groups for each genotype. p-value significance level was ****≤ 0.0001, ***≤ 0.001, ** ≤0.01, *≤ 0.05, ns > 0.05.

For UVB, heat, wound and salt, comparison for traits of control plants between WT and mutants were made to find out potential significant variation between these genotypes. In order to check whether certain abiotic treatment have more profound effect on the mutants in comparison to WT, significance test were performed between the untreated and treated plants of respective genotypes.

For ambient temperature, one-way ANOVA followed by Dunnett test was performed to determine the level of significance between the mean of control and treated plants using Prism software. p-value significance level used was ****≤ 0.0001, ***≤ 0.001, ** ≤0.01, *≤ 0.05, ns > 0.05.

7.4 Results

7.4.1 Analysis of gene expression atlas data of AtCDK8 and AtCYCC1 for abiotic stresses

In order to determine what kind of abiotic stresses regulate the expression of *CDK8* and *CYCC1* genes in *A. thaliana,* the gene expression atlas data was obtained from publicly available AtGenExpress database (http://jsp.weigelworld.org/expviz/expviz.jsp) (Figure 7.1 A-D, Table A4.1). As was observed, overall expression of *AtCYCC1.1* and *AtCYCC1.2* varied more widely between control and salt treated plants compared to *AtCDK8*. Besides, *AtCDK8* and *AtCYCC1.2* showed higher alteration in overall expression between untreated and UVB treated plants in

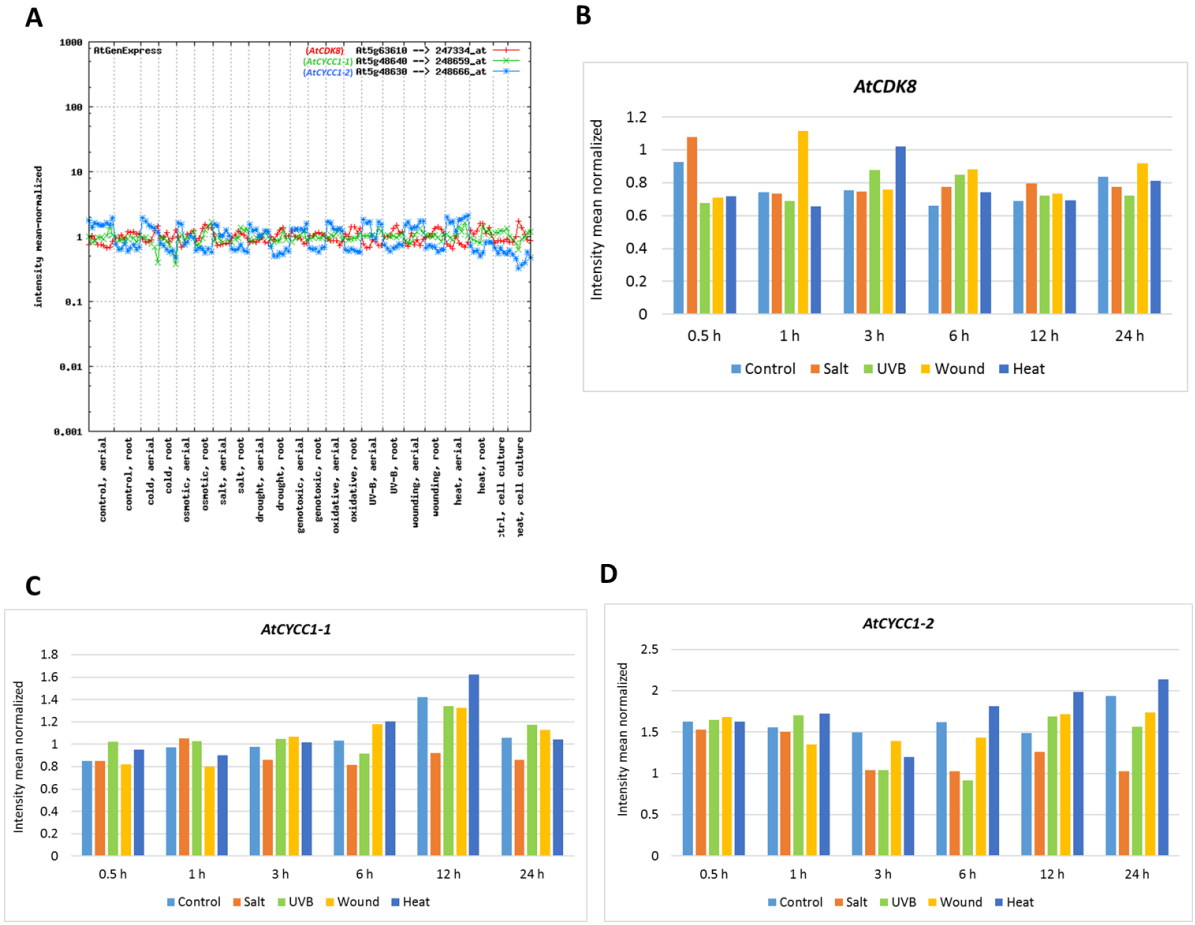


Figure 7.1. Expression atlas for *CDK8*, *CYCC1.1* and *CYCC1.2* gene in *A. thaliana* under various abiotic conditions. **(A)** Expression atlas under all conditions for *AtCDK8*, *AtCYCC1-1* and *AtCYCC1-2.* **(B-D)** Comparison of expression atlas of *AtCDK8*, *AtCYCC1-1* and *AtCYCC1-2* under various salt, UVB, wound and heat treatments. See appendix table A4.1 for various treatments used for these experiments.

comparison to *AtCYCC1.1.* A more or less similar scenario was observed for the expression of these three genes between control and heat treated plants. Moreover, *AtCDK8* showed increased alteration in overall expression relative to *AtCYCC1.1* and *AtCYCC1.2* between untreated and wound treated plants. Based on these findings as well as precedence of other mediator genes' involvement in regulating abiotic stress, it was decided to check whether *late3* and *late4* mutants regulate response specifically to the aforementioned factors.

7.4.2 late3 and late4 mutants do not regulate photomorphogenesis

Light is very important for post-seed germination growth in plants. In order to reveal the effect of *late3* and *late4* mutations on photomorphogenesis, leaf area and internode length were measured from respective mutants grown under continuous white, red, far-red, blue light as well as darkness.

Under darkness, *late3-1*, *late3-2* and *late4-1* mutants exhibited pronounced etiolated growth as they had significantly longer internodes relative to WT (Figure 7.2 A). There was no significant variation for this trait between WT and all the three mutants under red light. In contrast, under white and far-red lights, significant difference with WT for internode length occurred only for *late4-1* mutants, but not for the two alleles of *late3* mutation. Lastly, the two stronger alleles out of three from *late3* and *late4* mutations, i.e., *late3-2* and *late4-1* showed de-etiolated growth in terms of internode length as they had significantly longer internodes compared to WT under blue light.

As far as leaf area was concerned, the stronger *late3* allele, i.e., *late3-2* grew significantly smaller leaf compared to WT under both darkness and continuous white light (Figure 7.2 B). There was no significant difference for this trait for any of the mutants in comparison to WT under blue light. On the other hand, only the *late3* mutant alleles, but not *late4* developed significantly smaller leaf compared to WT under both red and far-red lights.

Since the mutants showed response towards both seedling etiolation and de-etiolation, therefore it seems that *LATE3* and *LATE4* are unlikely to be involved in controlling photomorphogenic development. However, irrespective of photomorphogenesis, *LATE3* is probably important for internode development under darkness and blue light and for leaf enlargement under all conditions apart from blue light. In contrast, *LATE4* may be vital for

internode development under all conditions apart from white light while it is likely to play no role for leaf expansion under both darkness and various light treatments.

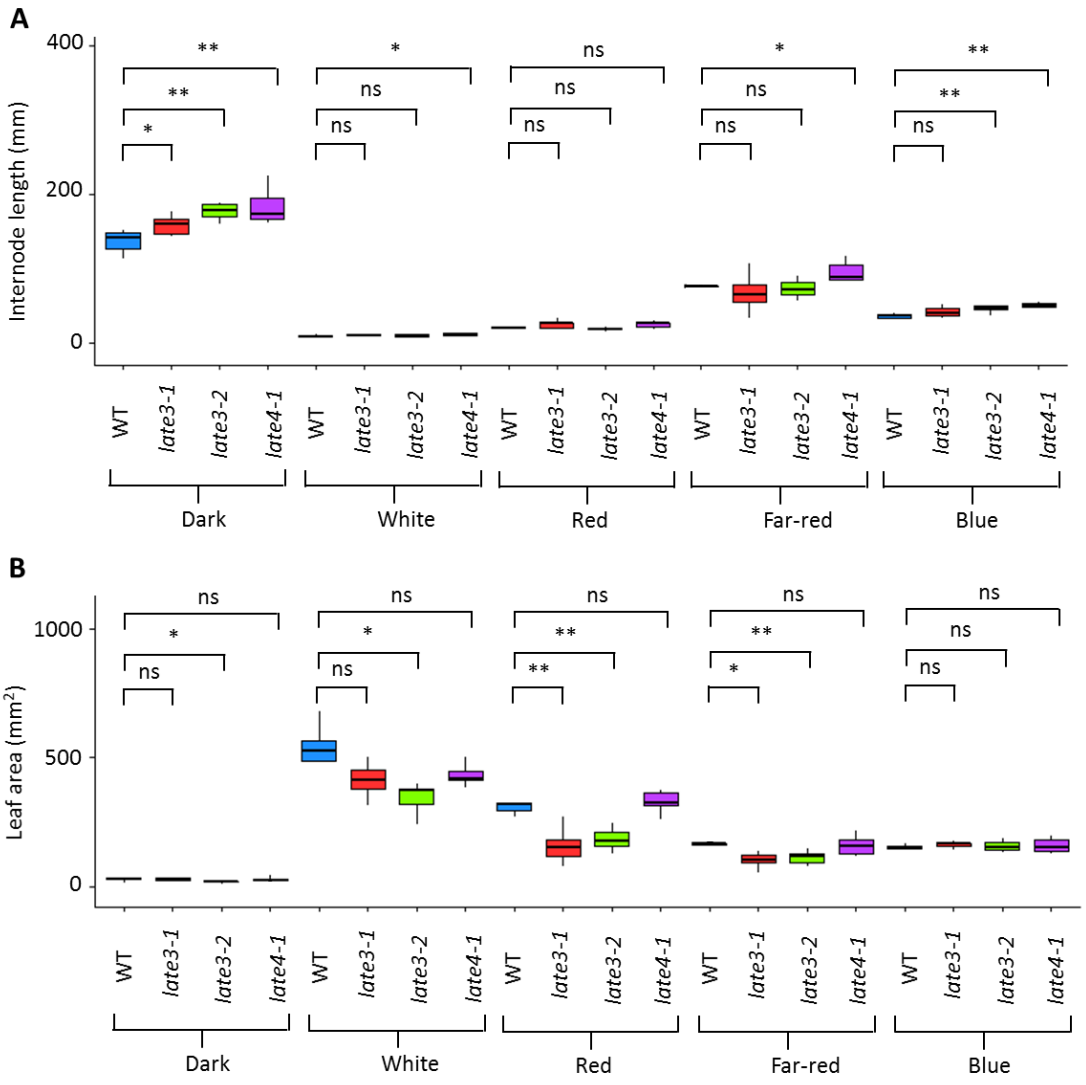


Figure 7.2. Responsiveness of WT (NGB5839), *late3* and *late4* mutants to various light conditions. Effect of different lights on **(A)** internode length and **(B)** leaf area. Internode length was measured as the length between node 1 and node 3. Leaf area was calculated as the product of length and width of a single leaflet at node 3. Box-whisker plots were generated from n = 7 plants. The Upper whisker represents highest 25% of the values whereas lower whisker indicates lower 25% of the values. The box represents middle 50% of the values falling within interquartile range (Q1 to Q3) and the line within the box shows median value. One-way ANOVA followed by Tukey test was performed to determine the level of significance between the median of wild-type (WT) and mutant plants for each condition. p-value significance level ***≤ 0.001, ** ≤0.01, *≤ 0.05, ns > 0.05.

7.4.3 late3 and late4 mutants are amenable to ambient temperature mediated flowering initiation in pea

Ambient temperature is important for determining flowering initiation in *A. thaliana* (Wigge 2013). In a previous experiment, pea cv. Alaska was grown under wide range of temperature conditions from 5° C to 24° C where days to first flowering reduced consistently from around 90 days at former condition to around 20 days at latter condition (Summerfield and Roberts 1988). In another study, cv. Alaska was exposed to 20/15 and 30/25° C day/night temperature regimes seven days after floral initiation for a week followed by transferring all the plants to 20/15° C (McDonald and Paulsen 1997). This resulted in atrophy of apical meristem and flowers to abort at 30/25° C compared to 20/15° C. Since, ambient temperature within this range shown to affect floral development, therefore 15° C, 20° C and 25° C temperatures were used in the present study for checking response of *late3* and *late4* mutants.

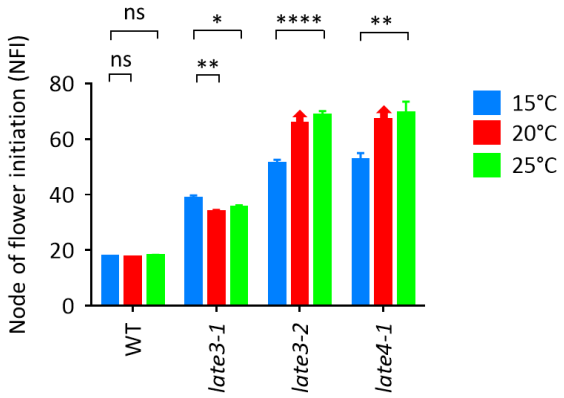


Figure 7.3. Responsiveness of WT (NGB5839), *late3* and *late4* mutants to ambient temperature. Effect of various ambient temperature on flowering initiation. *late3-2* and *late4-1* mutants did not flower at 20° C and graphs show the total number of nodes at which these mutants underwent senescence; for all other temperatures of each of the genotypes the graph is showing the node at which flowering initiation occurred. Data represents mean ± SE for n =4-6 plants. One-way ANOVA followed by Dunnett test was performed to determine the level of significance between the mean of control and treated plants using Prism software. p-value significance level **** ≤ 0.0001, ***≤ 0.001, ** ≤0.01, *≤ 0.05, ns > 0.05.

The WT flowered more or less at same node under all three temperature conditions (Figure 7.3). In contrast, weaker *late3-1* mutant showed significant variation in floral initiation at 20° C and 25° C in comparison to 15° C. Similarly, significant variation in commencement of

flowering was observed for stronger *late3-2* and *late4-1* mutants at 25° C in comparison to 15° C. A point to notice was that both *late3-2* and *late4-1* mutants did not flower at 20° C until around 65 nodes when these plants finally senesced naturally (Figure 7.3). Similar observation was made for *late4-1* mutants from the gene expression experiment discussed in chapter 6. However, as discussed in chapter 3, both these mutants flowered when they were grown under mostly natural light conditions in the glasshouse (appendix Figure A4.1). The lights that were used in the growth chamber lacks far-red spectrum. Therefore, these results suggest that *late3* and *late4* mutation affect ambient temperature coupled with light quality mediated flowering process in pea.

7.4.4 late3 and late4 mutations affect growth upon UVB exposure

UVB induced morphological changes include effect on leaf area, petiole length, stem elongation and shoot biomass (Surabhi et al. 2009; Bernal et al. 2015; Randriamanana et al. 2015; Robson et al. 2015). In order to test whether *late3* and *late4* mutations affect the response to UVB exposure, these mutants were subjected to UVB radition for different periods of time and various growth related traits were measured. At first, a pilot experiment was conducted in order to determine the level of UVB exposure that could cause significant damage in the WT genotype so that the same level or above treatment could be used for investigation involving the mutants. It was found that UVB exposure of 15 minutes was sufficient for significant change of only stem dry weight, but not for leaf fresh and dry weight, stem fresh weight and stem length (Figure 7.4 A-E). On the other hand, 30 minutes of exposure significantly reduced all the traits except stem dry weight.

Since 30 minutes of UVB exposure was found to be enough to cause significant damage in the WT, so a second round of experimentation was then carried out to assess the effect of *late3* and *late4* mutations by subjecting both WT and mutants to 40 and 80 minutes of UVB irradiance. In this case, it was observed that *late3-2* and *late4-1* mutants showed similar effect like WT for internode length after 40 and 80 minutes of UVB exposure. Only *late4-1* developed significantly smaller leaves after 80 minutes of UVB treatment (Figure 7.5 A-B). In contrast, both *late3-2* and *late4-1* mutants demonstrated significant reduction for petiole length and shoot dry weight in UVB treated plants compared to control plants (Figure 7.5 C-D). These results suggest that *LATE3* and *LATE4* genes are likely to be directly or indirectly involved in the regulatory process that mediate plant response to UVB stress.

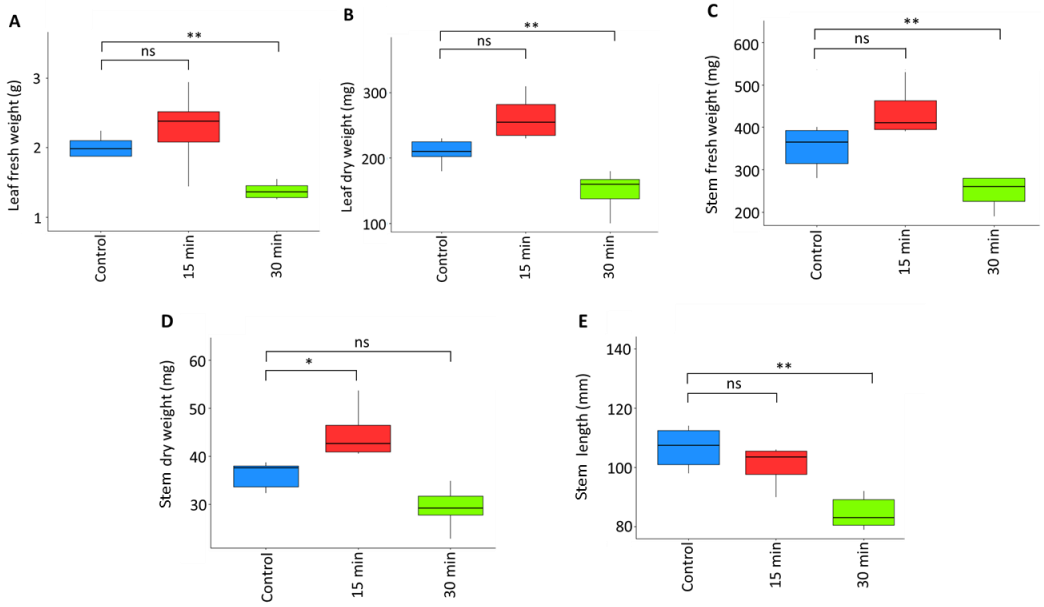


Figure 7.4. Responsiveness of WT (NGB5839) pea plants to UVB stress. Effect of UVB stress on **(A)** leaf fresh weight, **(B)** leaf dry weight, **(C)** stem fresh weight, **(D)** stem dry weight and **(E)** stem length. Stem length was measured as the length from cotyledon to bottom of shoot apex. Samples were dried in an oven at 70° C for 96 hours for measuring the dry weight. Box-whisker plots were generated from n = 6 plants. The Upper wisker represents highest 25% of the values whereas lower wishker indicates lower 25% of the values. The box represents middle 50% of the values falling within interquartile range (Q1 to Q3) and the line within the box shows median value. One-way ANOVA followed by Tukey test was performed to determine the level of significance between the median of control and treated plants. p-value significance level ***≤ 0.001, ** ≤0.01, *≤ 0.05, ns > 0.05.

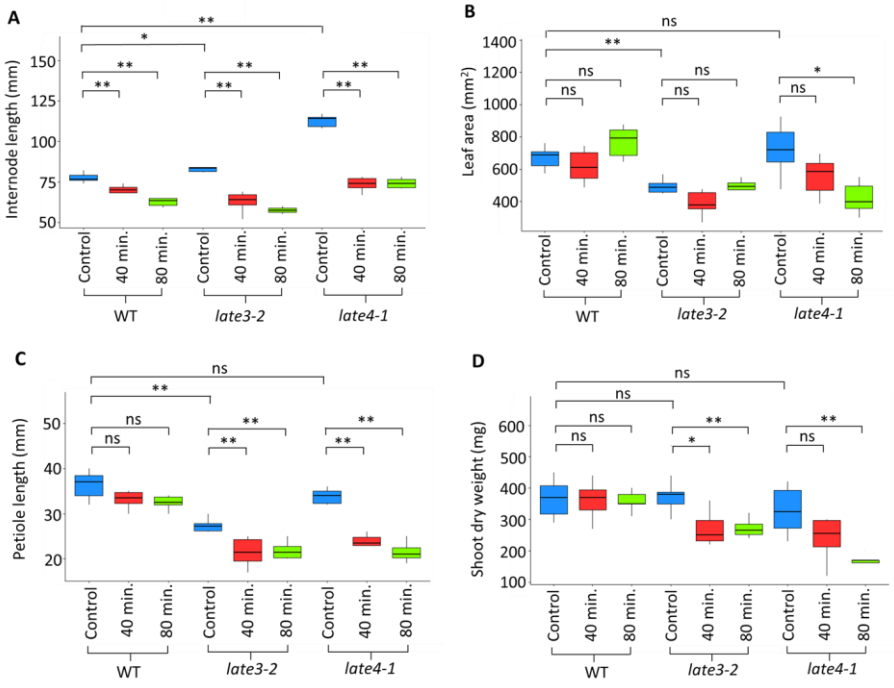


Figure 7.5. Responsiveness of WT, *late3* and *late4* mutants to UVB stress. **(continued next page)**

Figure 7.5. (continued) Effect of UVB on **(A)** internode length, **(B)** leaf area, **(C)** petiole length and **(D)** shoot dry weight. Internode length was measured as the length between node 1 and node 8. Leaf area was calculated as the product of length and width of a single leaflet from leaf 8 while petiole length illustrates the length of petiole from leaf 8. Samples were dried in oven at 70° C for 96 hours for measuring the dry weight. Box-whisker plots were generated from n = 6 plants. The Upper whisker represents highest 25% of the values whereas lower whisker indicates lower 25% of the values. The box represents middle 50% of the values falling within interquartile range (Q1 to Q3) and the line within the box shows median value. One-way ANOVA followed by Tukey test was performed to determine the level of significance between the median of control and treated plants for each genotype. p-value significance level ***≤ 0.001, ** ≤0.01, *≤ 0.05, ns > 0.05.

7.4.5 late3 and late4 mutants show increased sensitivity to heat stress

Phenotypic symptoms of heat stress include reduction in plant height, leaf area, total biomass etc. (Hasanuzzaman et al. 2013; Zandalinas et al. 2016; Giri 2013; Liu et al. 2017; Sehgal et al. 2017). In a previous experiment, exposure to 40° C for 2.5 minutes was found to highly affect photosynthetic system such as thylakoid activity in pea (McDonald and Paulsen 1997). Therefore, the initial pilot experiment involving WT genotype was designed by exposing the plants to a nearly similar temperature, i.e., 38° C for 1 to 4 hours. It was found that even 4

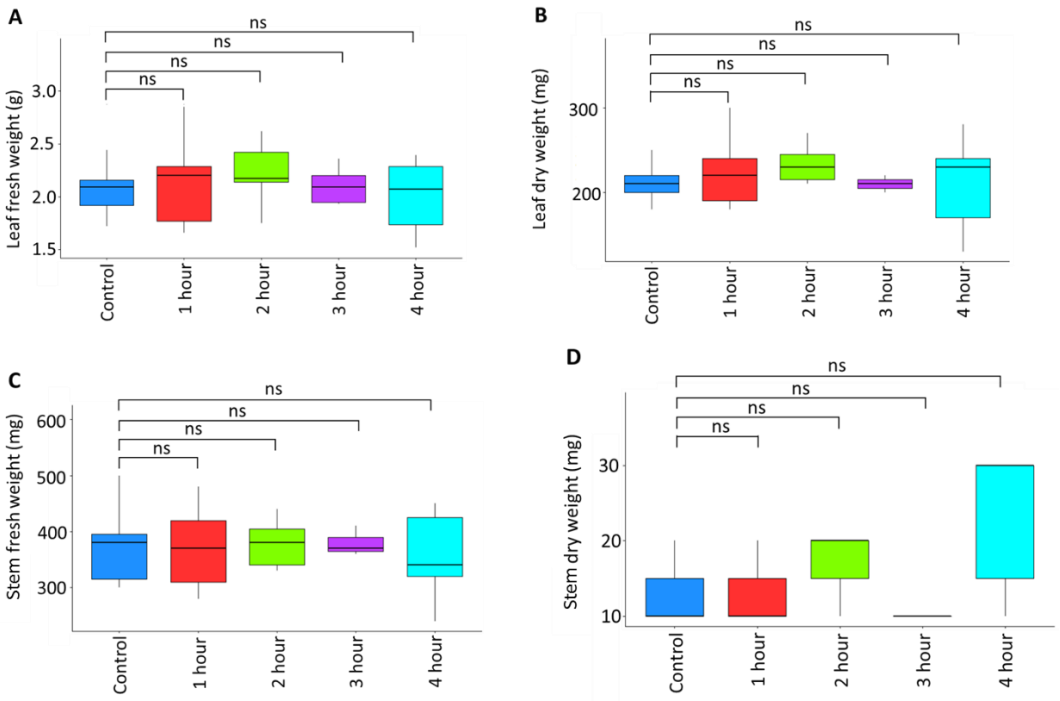


Figure 7.6. Responsiveness of WT (NGB5939) pea plants to 38° C heat stress. Effect of 38° C heat stress on **(A)** leaf fresh weight, **(B)** leaf dry weight, **(C)** stem fresh weight and **(D)** stem dry weight. Samples were dried in oven at 70° C for 96 hours for measuring the dry weight. **(continued next page)**

Figure 7.6. (continued) Box-whisker plots were generated from n = 7 plants. The Upper whisker represents highest 25% of the values whereas lower whisker indicates lower 25% of the values. The box represents middle 50% of the values falling within interquartile range (Q1 to Q3) and the line within the box shows median value. One-way ANOVA followed by Tukey test was performed to determine the level of significance between the median of control and treated plants. p-value significance level ***≤ 0.001, ** ≤0.01, *≤ 0.05, ns > 0.05.

hours of treatment was not enough to cause significant effect for traits such as fresh and dry weight of leaf and stem (Figure 7.6 A-D).

Based on results of the initial experiment, the next round of investigation using *late3-2* and *late4-1* mutants treated the plants with more severe heat stress, i.e., 45° C for 5 and 8 hours.

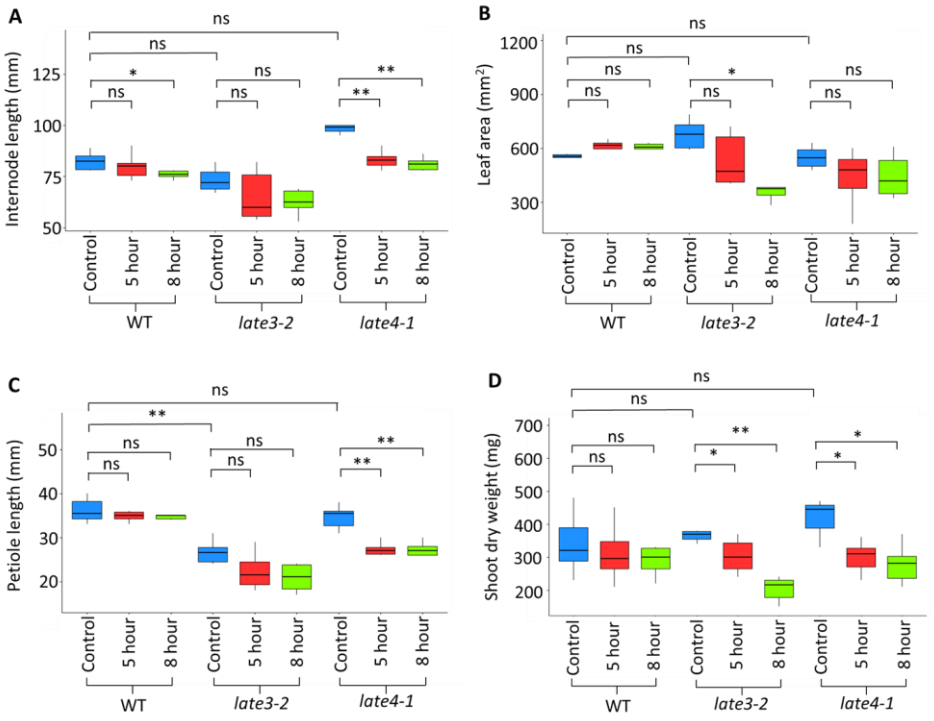


Figure 7.7. Responsiveness of WT (NGB5839), *late3* and *late4* mutants to 45° C heat stress. Effect of 45° C heat stress on **(A)** internode length, **(B)** leaf area, **(C)** petiole length and **(D)** shoot dry weight. Internode length was measured as the length between node 1 and node 8. Leaf area was calculated as the product of length and width of a single leaflet from leaf 8 while petiole length illustrates the length of petiole from leaf 8. Samples were dried in oven at 70° C for 96 hours for measuring the dry weight. Box-whisker plots were generated from n = 6 plants. The Upper whisker represents highest 25% of the values whereas lower whishker indicates lower 25% of the values. The box represents middle 50% of the values falling within interquartile range (Q1 to Q3) and the line within the box shows median value. One-way ANOVA followed by Tukey test was performed to determine the level of significance between the median of control and treated plants for each genotype. p-value significance level ***≤ 0.001, ** ≤0.01, *≤ 0.05, ns > 0.05.

As was observed, the WT still did not show significant effect for most of the traits e.g., leaf area, petiole length and shoot dry weight after 8 hours of treatment (Figure 7.7 A-D). But, significant difference was observed only for internode length at 8 hour time point. In contrast, *late3-2* mutants showed significant reduction in leaf area for 8 hour and shoot dry weight at both time points, but not for internode and petiole length (Figure 7.7 A-D). As far as *late4-1* was concerned, significant decrease was found for internode length, petiole length and shoot dry weight after both 5 and 8 hours of heat stress. In contrast, there was no significant change for leaf area. Altogether, the results from this experiment indicates that WT pea is quite heat tolerant and *LATE3* and *LATE4* are vital genes for regulating response towards heat stress in pea.

7.4.6 late3 and late4 mutants influence reponse to wounding

The gene atlast database of *A. thaliana* revealed that wounding regulates expression of *AtCDK8* and *AtCYCC1* genes (Figure 7.1 A-D). Since plants exhibit local and system response upon wound stress similar to biotic stress treatment up on leaf (Savatin et al. 2014; León et al. 2001), so alteration in phenotypic traits such as root/shoot biomass (Lewandowski et al. 2013; Llugany et al. 2013) was studied in the current study. Initially, the WT genotype was subjected to two different treatments using a thumbtack: treatment 1 - 10 holes on leaf 3 (5

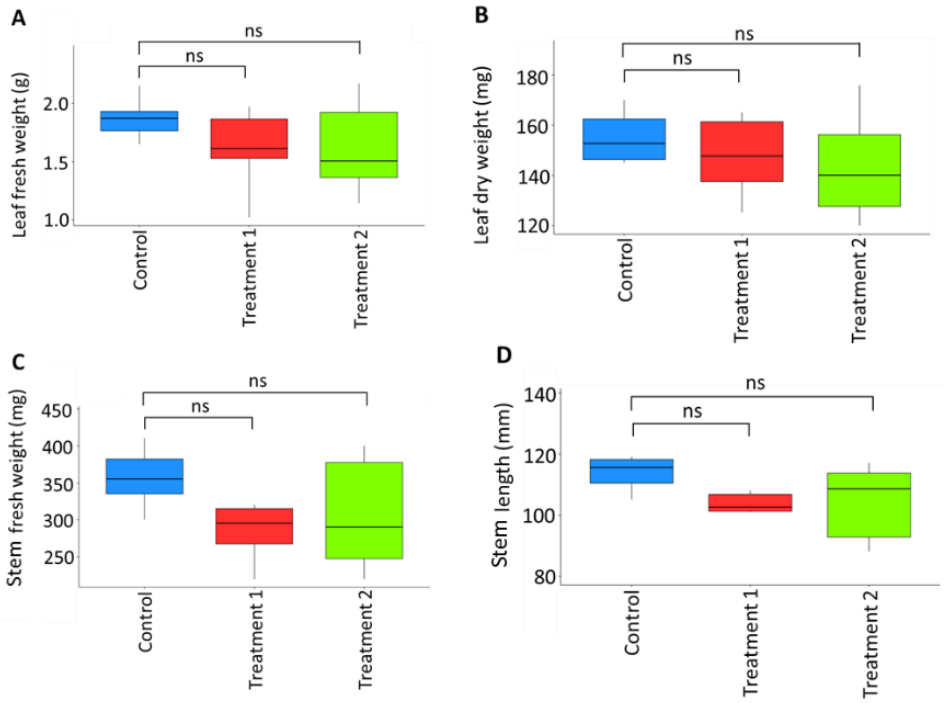


Figure 7.8. Responsiveness of WT (NGB5939) pea plants to wound stress. **(continued next page)**

Figure 7.8. (continued) Effect of wound stress on **(A)** leaf fresh weight, **(B)** leaf dry weight, **(C)** stem fresh weight and **(D)** stem length. Stem length was measured as the length from cotyledon to shoot apex. Samples were dried in oven at 70° C for 96 hours for measuring the dry weight. Box-whisker plots were generated from n = 6 plants. The Upper wisker represents highest 25% of the values whereas lower whisker indicates lower 25% of the values. The box represents middle 50% of the values falling within interquartile range (Q1 to Q3) and the line within the box shows median value. One-way ANOVA followed by Tukey test was performed to determine the level of significance between the median of control and treated plants. p-value significance level ***≤ 0.001, ** ≤0.01, *≤ 0.05, ns > 0.05.

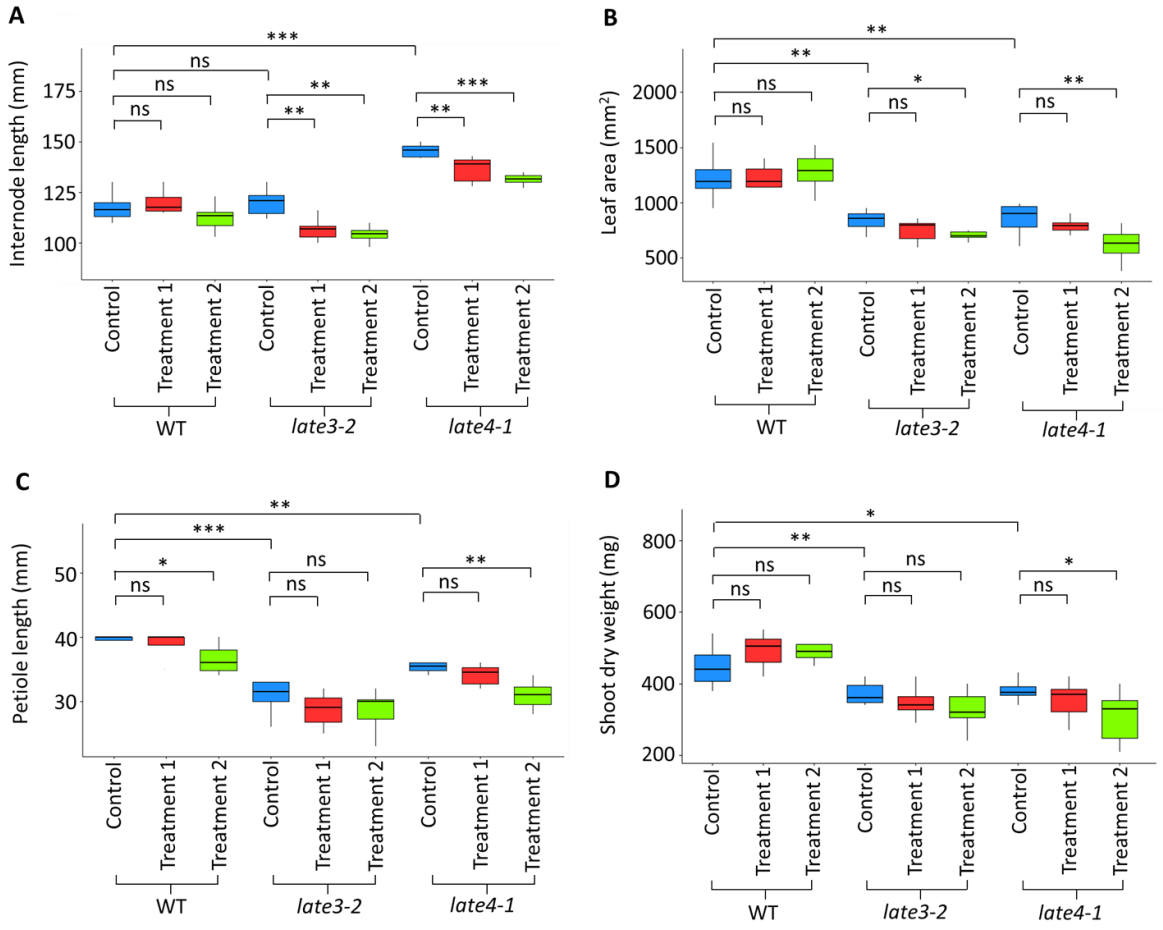


Figure 7.9. Responsiveness of WT (NGB5839), *late3* and *late4* mutants to wound stress. Effect of wound stress on **(A)** internode length, **(B)** leaf area, **(C)** petiole length and **(D)** shoot dry weight. Data from figure 3.11 has been re-plotted for shoot dry weight in control plants. Internode length was measured as the length between node 1 and node 10. Leaf area was calculated as the product of length and width of a single leaflet from leaf 8 while petiole length illustrates the length of petiole from leaf 8. Samples were dried in oven at 70° C for 96 hours for measuring the dry weight. Box-whisker plots were generated from n = 8 plants. The Upper whisker represents highest 25% of the values whereas lower whisker indicates lower 25% of the values. The box represents middle 50% of the values falling within interquartile range (Q1 to Q3) and the line within the box shows median value. One-way ANOVA followed by Tukey test was performed to determine the level of significance between the median of control and treated plants for each genotype. p-value significance level ***≤ 0.001, ** ≤0.01, *≤ 0.05, ns > 0.05.

holes/leaflet), treatment 2 - 10 holes on leaf 2 and 3 (5 holes/leaflet). No major effect of both these treatments was observed in the WT genotype for all the phenotpes, i.e., leaf fresh and dry weight, stem fresh weight and stem length (Figure 7.8 A-D). These results suggested that higher doses of wound treatment would be needed for the next set of experiments involving *late3* and *late4* genotypes.

In the second round of experiment, WT, *late3-2* and *late4-1* genotypes were given two treatments using again the thumbtack. These treatments were as follows: treatment 1 - 40 holes on leaf 2, 3 (20 holes/leaflet); treatment 2 - 80 holes on leaf 2, 3 (40 holes/leaflet). Here, the WT did not exhibit any substantial effect of wound stress for all the phenotypes such as internode length, leaf area, and shoot dry weight except petiole length at treatment 2 (Figure 7.9 A-D). On the other hand, responsiveness of *late3-2* was significant for internode length and leaf area whereas the opposite scenario was observed for petiole length and shoot dry weight. For *late4-1*, significant reduction was observed for all the traits between the control and treated plants mostly after treatment 2. Based on the findings of this experiment, it can be stated that mutation at both *LATE3* and *LATE4* loci affect growth in pea upon wounding.

7.4.7 late3 and late4 mutants increase susceptibility to salt stress

From the gene atlas database of *A. thaliana*, it was found that expression of *AtCDK8* and *AtCYCC1* genes are affected by salt stress. So, the WT genotype was initially treated with various salt concentrations. Salt stress is known to affect phenotypic traits such as root length, fresh and dry weight of root and shoot (Li et al. 2013a; Ventura et al. 2014; Duan et al. 2015; Robin et al. 2016; Fu et al. 2017). In a previous experiment, 70mM NaCl concentration was found to affect growth of a salt sensitive pea variety (cv. Challis) (Hernandez et al. 1995). Therefore, three different salt treatments consisting of 50 mM, 75 mM and 100 mM NaCl in 500 ml total volume of 1% nutrient solution were applied to the treatment plants on day 14 in order to find out the level of salt tolerance/sensitivity of WT NGB5839 (cv. Torsdag background) genotype. Control plants were applied 500 ml of 1% nutrient solution only. In this study, phenotypic traits measured after two more weeks revealed that considerable positive growth occurred under all three salt treatments as significant increase in leaf fresh and dry weight, stem fresh and dry weight and stem length was found in the treated plants compared to the untreated ones (Figure 7.10 A-E). Further discussion on this aspect would be carried out in section 7.5.5. In a simultaneous experiment at the School of Natural Sciences,

University of Tasmania carried out by Dr. Eloise Foo (personal communication) it was found that cv. Torsdag genotype remained more or less healthy even after 200 mM of NaCl application. These findings indicated that even higher concentrations of salt treatment need to be applied for finding negative effect of salt stress upon the WT, *late3* and *late4* mutants.

Next, WT, *late3-2* and *late4-1* genotypes were treated with 250 mM and 500 mM of NaCl in total volume of 100 ml water from day 6 to 11. For all three genotypes, no significant variation was found for primary root length, internode length and shoot dry weight at 250 mM NaCl (Figure 7.11 A-C). But, significant reduction was observed for root dry weight in all the three

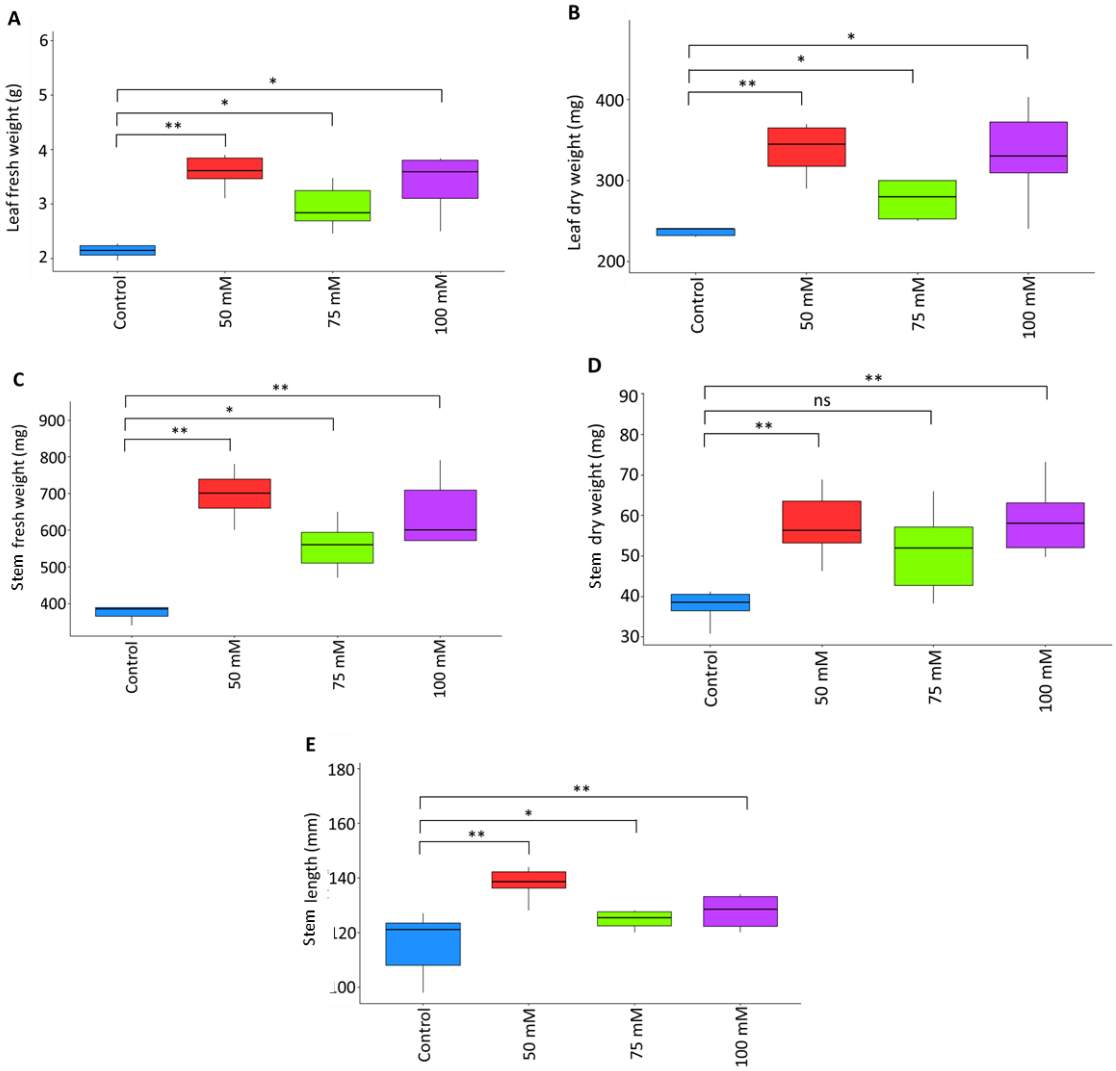


Figure 7.10. Responsiveness of WT (NGB5839) pea plants to salt stress. Effect of salt stress on **(A)** leaf fresh weight, **(B)** leaf dry weight, **(C)** stem fresh weight, **(D)** stem dry weight and **(E)** stem length. Stem length was measured as the length from cotyledon to shoot apex. Samples were dried in oven at 70° C for 96 hours for measuring the dry weight. Box-whisker plots were generated from n = 6 plants. **(continued next page)**

Figure 7.10. (continued) The Upper whisker represents highest 25% of the values whereas lower whisker indicates lower 25% of the values. The box represents middle 50% of the values falling within interquartile range (Q1 to Q3) and the line within the box shows median value. One-way ANOVA followed by Tukey test was performed to determine the level of significance between the median of control and treated plants. p-value significance level ***≤ 0.001, ** ≤0.01, *≤ 0.05, ns > 0.05.

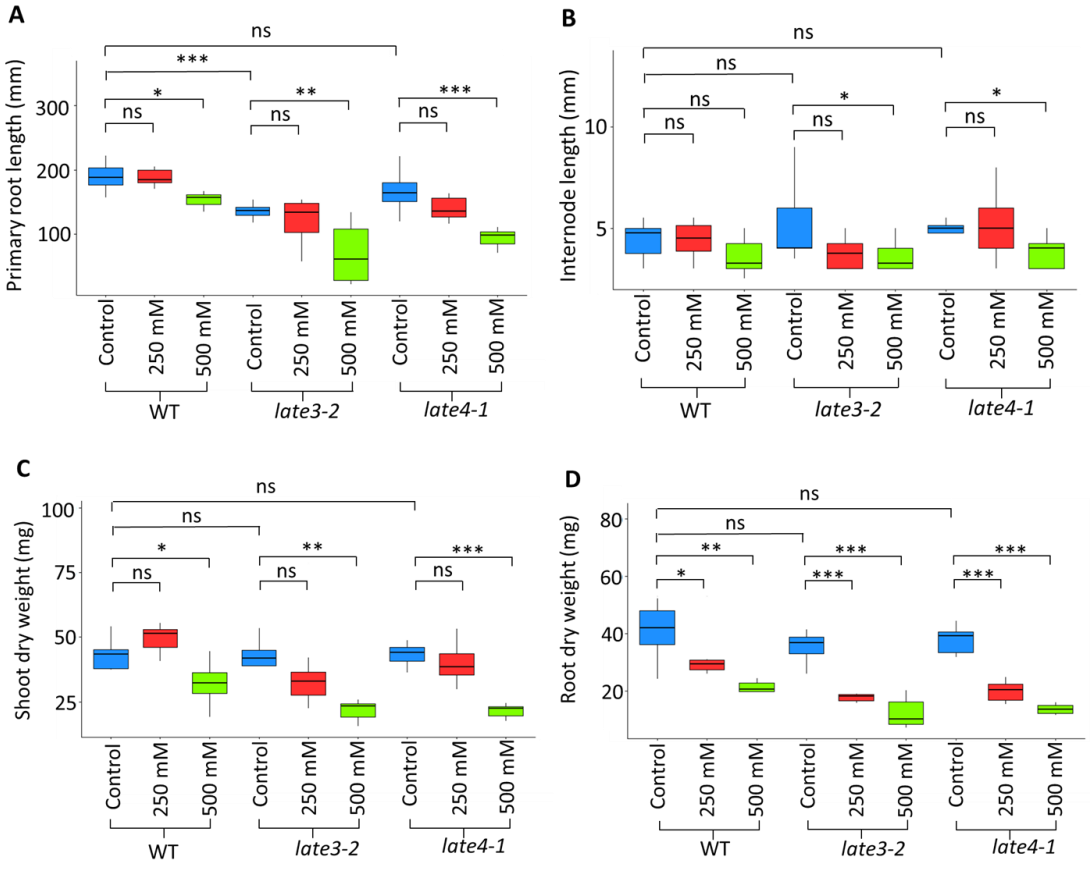


Figure 7.11. Responsiveness of WT (NGB5839), *late3* and *late4* mutants to salt stress. Effect of salt stress on **(A)** primary root length, **(B)** internode length, **(C)** shoot dry weight and **(D)** root dry weight under. Data from figure 3.11 has been re-plotted for primary root length and root dry weight in control plants. Primary root length was measured as the length from cotyledon to the tip of primary root. Internode length was measured as the length between node 1 and node 2. Samples were dried in oven at 70° C for 96 hours for measuring the dry weight. Box-whisker plots were generated from n = 8 plants. The Upper whisker represents highest 25% of the values whereas lower whishker indicates lower 25% of the values. The box represents middle 50% of the values falling within interquartile range (Q1 to Q3) and the line within the box shows median value. One-way ANOVA followed by Tukey test was performed to determine the level of significance between the median of control and treated plants for each genotype. p-value significance level ***≤ 0.001, ** ≤0.01, *≤ 0.05, ns > 0.05.

genotypes for this concentration (Figure 7.11 D). At 500 mM NaCl, no significant alteration in WT occurred for internode length whereas contrasting situation was visible for *late3-2* and *late4-1* mutants (Figure 7.11 B). In contrast, primary root length, shoot and root dry weight were significantly decreased in all the three genotypes at the same concentration (Figure 7.11 A, C, D). However, the magnitude of reduction of these traits in *late3-2* and *late4-1* seemed be higher than what was observed in WT at 500 mM NaCl between treated and untreated plants. For instance, primary root length was reduced by 17%, 48% and 43% in WT, *late3-2* and *late4-1* respectively. Similarly, there was 22%, 46% and 50% decrease in WT, *late3-2* and *late4-1* respectively for shoot dry weight. Likewise, root dry weight fell off by 50%, 64% and 64% in WT, *late3-2* and *late4-1* respectively. Altogether, these results suggest that *late3* and *late4* mutants exhibit moderate response upon salt stress.

7.5 Discussion

7.5.1 LATE3 and LATE4 seems to be important for general dark and light mediated growth, but not specifically for photomorphogenic development

Light quality is an important regulator of early stage seedling germination and development in plants. In the present study, it was observed that *late3* mutants affect seedling etiolation and de-etiolation for internode length under blue light and for leaf area under white, red and far-red light (Figure 7.2 A-B). Similarly, *late4-1* mutants seemed to influence seedling etiolation and de-etiolation for internode only under white, far-red and blue light conditions (Figure 7.2 A-B). The observed pattern of difference in response towards a certain treatment between *late3* and *late4* mutants could give hints towards existence of mediator independent *CDK8* function which is known in human (Firestein and Hahn 2009) and it might also be due to the use of a particular light irradiance. Usually, the existence of light and dark specific effect identifies a mutant as photomorphogenic. Some examples of such photomorphogenic pea mutants are *fun*, *lv* and *lip1* mutants where the underlying genes were identified as pea homologues of *PHYA*, *PHYB* and *COP1* (Frances et al. 1992; Murfet 1993; Weller and Reid 1993; Weller et al. 1997a; Sullivan and Gray 2000; Weller et al. 2001; Weller et al. 2004). The results from the present study suggested that both *LATE3* and *LATE4* may have general effects on growth that are independent of the light/dark conditions and therefore are not related to photomorphogenesis.

In a recent study involving differential expression of the mediator genes in *A. thaliana* seedlings, it was found that darkness downregulates all the mediator genes apart from *MED15* and *MED37* whereas light upregulates all the mediator genes (Pasrija and Thakur 2012; Dwivedi et al. 2017). Based on these findings and perceiving the existence of a conserved process in pea, the results regarding *LATE3/PsCDK8* and *LATE4/PsCYCC1* genes obtained during the current study was in agreement with that of *CDK8* and *CYCC1* genes of *A. thaliana*. In chapter 6, the results from AraNET v2.0 predicted that *MtCDK8* and *MtCYCC1* are likely to act in regulation of response to light stimulus (Table 6.2, 6.3). Besides, findings from PlantTFDB v4.0 predicted that bZIP and Dof transcription factors having known role in light signalling pathway are likely to regulate function of *CDK8* and *CYCC1* genes in *M. truncatula* as well as in pea (Table 6.4, 6.5). Therefore, it can be hypothesized function of *LATE3* and

LATE4 might be regulated by light where the aforementioned TFs act in perceiving the relevant light signal. Therefore, future experiments in this direction would be of high interest.

7.5.2 LATE3 and LATE4 are likely to be involved in the regulatory process that control ambient temperature dependent flowering

In the present study, responsiveness of *late3* and *late4* mutants towards three different ambient temperature regimes, i.e., 15° C, 20° C and 25° C were checked (Figure 7.3). The results revealed substantial response of both the mutants for ambient temperature mediated thermomorphogenesis process. Under low ambient temperature during LD condition, *A. thaliana* MADS box transcription factors SVP, FLM-β and FLC form a suppressor complex that block activity of floral promoting genes such as *SOC1* and *FT* (Capovilla et al. 2015). Besides, miR156 contributes to repression of *FT*. Under higher temperature, a dominant version of FLM namely FLM-α block the repressive complex formed by MADS box transcription factor which leads to promotion of *FT* activity. Previously, three *SVP* like genes namely *SVPa*, *SVPb* and *SVPc* were identified in legumes including pea (Sussmilch 2014). Among them, *PsSVPc* expression was found to be negatively regulated by *VEG2*/*PsFD*. In the present study, it was found that *LATE3* and *LATE4* positively regulate expression of *FTa*, *FTc* and *VEG2*/*PsFD* (until day 28, section 6.4.5). Therefore, it can be speculated that misregulation of *PsSVPc* may have contributed to marked delay in flowering in higher ambient temperature (25° C) compared to low ambient temperature (15° C) in the stronger alleles *late3-2* and *late4-1* by acting negatively upon pea *FT* genes in a manner mentioned earlier for *A. thaliana*. Future expression studies of *PsSVPc* and also *PsSVPa*, *PsSVPb* as well as interaction studies of these genes in combination with *FT*s and *VEG2* in *late3* and *late4* mutants is likely to reveal better understanding in this regard.

Another interesting finding of the present study was that *late3-2* and *late4-1* mutants did not flower at 20° C in the growth chamber light conditions compared to that of glass house (Figure 7.3). The weaker *late3-1* flowered considerably later in the growth chamber compared to glasshouse whereas the WT flowered more or less at around the same time (appendix Figure A4.1). A key difference between the two conditions is that there is no far-red light in the growth chamber and the intensity of light is probably less. It is known in *A. thaliana* that phyB speeds up the process of flowering by triggering shade avoidance syndrome (SAS) under low R:FR light condition and phyB acts through JA in this process (Cerdan and Chory 2003; Iñigo

et al. 2012; Moreno et al. 2009; Izaguirre et al. 2006). Besides, JA signalling has known role in light responses (Riemann et al. 2008; Riemann et al. 2009; Robson et al. 2010). *AtMED25* was found to act downstream of phyB to regulate flowering and also to affect JA responsive genes in a temperature dependent manner which led to suggestion that *AtMED25* integrates various environmental signals that are known to interact with light quality mediated signalling (Kidd et al. 2009; Iñigo et al. 2012). Therefore, the JA oriented signalling process described for *AtMED25* involving light quality and temperature may be evolutionary conserved in pea and *LATE3*/*PsCDK8* and *LATE4*/*PsCYCC1* which are part of the large mediator complex might play vital role in this direction at 20° C in the absence of far-red light. Further genetic interaction studies between *LATE3*/*LATE4* and *PsPHYB* would probably strengthen knowledge about this aspect.

Another explanation could be that difference in light intensity/quantity between the two growth conditions may have caused the substantial variation in flowering initiation in *late3* and *late4* mutants at 20° C. In a recent study involving *A. thaliana*, it was shown that high light condition generates a chloroplast derived signal mediated by a PHD transcription factor called PTM which then binds at the promoter region of *FLC* in the nucleus along with FVE (Feng et al. 2016; Susila et al. 2016). In this way, PTM-FVE complex suppress the function of *FLC* and accelerates flowering. This process might be evolutionary conserved in pea as well and *LATE3*/*PsCDK8* and *LATE4*/*PsCYCC1* might play a role in this direction. A third explanation could be that temperature in the glasshouse fluctuates and is not consistently 20° C and photoperiod also varies by two hours compared to growth chamber. However, since both the mutants flowered at both 15° C and 25° C, so it is unlikely that slight variation in photoperiod and temperature have caused such substantial variation. Taken together, all these findings provide hint that *LATE3*/*PsCDK8* and *LATE4*/*PsCYCC1* are crucial for regulating the dynamic process that determine development of pea through interaction with ambient temperature and light quality.

7.5.3 LATE3 and LATE4 are crucial for mediating response to UVB stress

During the present study, findings from UVB (280-315 nm) exposure of *late3* and *late4* mutants revealed considerable damage in these mutants (Figure 7.5). In *A. thaliana*, the photoreceptor *UVR8* perceives UVB light and leads the response of plant in this direction (Tilbrook et al. 2013; Heijde and Ulm 2012; Li et al. 2009). Once UVR8 is switched on, it then

interacts with COP1 which is under negative regulation of light activated phytochromes and crytochromes and this leads to stabilization of HY5 and downstream signal transduction activation. As a result, proteins involved in UV protection such as those from the phenylpropanoid pathway, reactive oxygen species scavenging, and DNA damage repair (e.g. phtolyases) are generated. In has already been reported that mutation in *AtMED5a* and *AtMED5b* causes upregulation of various phenylpropanoid genes producing significant amount of phenylpropanoids (Bonawitz et al. 2014). Therefore, it has been suggested that the *AtMED5* is likely to play a major role in plant defence upon UV radiation (Dwivedi et al. 2017). Besides, *AtMED16* was found to regulate response to UV-C stress (Wathugala et al. 2012). In addition, differential expression of *AtCDK8* and *AtCYCC1* upon UVB exposure was observed via the gene atlas database of *A. thaliana* (Figure 7.1, Table A4.1). In chapter 6, the results from AraNET v2.0 predicted that *MtCDK8* and *MtCYCC1* are likely to act in regulation of DNA endoduplication and reactive oxygen species metabolic process (Table 6.2, 6.3). Based on these information, it can be hypothesized that *LATE3*/*PCDK8* and *LATE4*/*PsCYCC1* genes' expression are probably regulated by UVB treatment which then activates the relevant response such as phenylpropanoid, DNA repair and ROS generation pathways in pea. Pea homologues of phytochromes phyA and phyB (Weller et al. 2001) and cryptomechromes cry1 (Platten et al. 2005) as well as divergent *AtHY5* (bZIP transcription factor) orthologue namely *LONG1* (Weller et al. 2009) have been identified already. Future functional relationship of these pea homologs with *LATE3* and *LATE4* would be of future interest for better understanding the underlying UVB induced defence response.

7.5.4 LATE3 and LATE4 are important for heat stress response

In the quest of understanding role of *LATE3* and *LATE4* in regulating response towards heat stress, the respective mutants were exposed to 45° C temperature. Findings from this investigation unravelled that mutation in both the genes significantly affect overall pea growth when exposed to high temperature stress (Figure 7.7). Role of some mediator subunits in relation to heat stress response is known. It was reported previously that *MED16* and *MED23* are important for heat-shock-specific gene expression in *Drosophila* (Kim et al. 2004). Likewise, it was shown by providing 39° C heat shock upon *S. cerevisiae* that *Heat Shock Factor 1 (HSF1)* recruits the CDK8 module at the regulatory region of *HSF1* - target genes independently of core mediator (Anandhakumar et al. 2016). In *A. thaliana*, defence against

heat stress was reported to be conferred by transgenic experimentation involving heat shock protein HSP70 from *Trichoderma harzianum* (Montero-Barrientos et al. 2010; Hasanuzzaman et al. 2013). Among the *A. thaliana* mediator complex components, *AtMED37* are annotated as HSP70 proteins and their orthologues in *M. truncutula* and pea have been mentioned in chapter 5. The HSP70 suppress function of key heat stress response gene *HsfA1* (Hahn et al. 2011; Yamada et al. 2007) which suggest likely important role of this mediator complex subunit in the heat stress response regulatory process in plants. In addition, it was found from the *A. thaliana* gene atlas that *AtCDK8* and *AtCYCC1* are differentially regulated by heat stress (Figure 7.1, Table A4.1). Moreover, the forecasted functions of *MtCDK8* and *MtCYCC1* presented in chapter 6 involves response to heat stress (Table 6.2, 6.3). Based on these information, a conserved mechanism can be speculated where *LATE3*/*PsCDK8* and *LATE4*/*PsCYCC1* along with predicted transcription factors shown in Table 6.4, 6.5 may have role in the regulatory process that deal with heat stress in pea.

7.5.5 LATE3 and LATE4 are important regulators of response towards wound and salt stress in pea

Prompt healing after an injury is very important for plants in order to achieve proper development. Investigations with different levels of wounding treatments unveiled that *late3* and *late4* mutants considerably affect response towards such mechanical stress (Figure 7.9). Studies involving *A. thaliana* have revealed that plant's response to mechanical injury is similar to that of pathogen attack. In this case, ROS such as H_2O_2 generation is important as part of local as well as systemic response (Mittler et al. 2011; Suzuki and Mittler 2012; Savatin et al. 2014). Besides, JA is considered to be vital for systemic response (Penacortes et al. 1995; Bergey et al. 1996; Creelman and Mullet 1997; Savatin et al. 2014). Likewise, ABA was also reported to act positively to promote programmed cell death in order to contain the detrimental effect of wound (Bostock and Stermer 1989; Savatin et al. 2014). It has been reported previously that *AtCDK8*/*AtCDKE1* is a regulator of H_2O_2 stress in *A. thaliana* (Ng et al. 2013). In addition, *AtMED25* was found to mediate JA and ABA signalling by direct physical interaction with bHLH transcription factor MYC2 and bZIP transcription factor, ABI5 (Chen et al. 2012). In the gene atlast of *A. thaliana*, it was found that wounding differentially regulates expression of *AtCDK8* and *AtCYCC1* (Figure 7.1, Table A4.1). Moreover, the predicted functions of *MtCDK8* and *MtCYCC1* revealed that they are likely to be involved in hormone

signalling and ROS metabolic processes. Taken together, it can therefore be hypothesized that *LATE3*/*PsCDK8* and *LATE4*/*PsCYCC1* act in a similar fashion to regulate the response towards wounding in pea where the predicted transcription factors presented in table 6.4 and 6.5 might also play important role. Future experimentation is recommended in this direction by checking expression of homologues of key pea JA and ABA signalling genes in *late3* and *late4* mutants under wound stress.

Salt stress is another environmental factor that negatively affects crop yield leading to economic losses around the world. In the first round of experiments involving application of low concentration of salt (50 mM, 75mM and 100 mM NaCl) unto WT genotype, it was found that the plant have positive growth rather than any reduction (Figure 7.10). Previously, it was found that a pea variety cv. Green Arrow can tolerate 70 mM NaCl and it also showed higher growth at 10 and 30 mM NaCl compared to untreated plants (Najafi et al. 2007). The reason behind positive growth could be that WT pea plants had minimized usage of K^+ in the presence of elevated Na^+ level up to a certain limit in order to carry out the processes such as turgor pressure maintenance, vacuole and cell expansion. K^+ and Na^+ are have similarities in chemical and structural properties and therefore they exert similar behavioural patterns (Benito et al. 2014). When Na^+ is available, then it can be used up to a certain degree and plant can exploit available K^+ for other more K^+ specific purposes. Indeed, growth enhancement and quality improvement by adding Na^+ to growth medium have been reported in research involving red beet and other crop species such as celery, carrots, forage crops (Subbarao et al. 1999; Subbarao et al. 2003; Marschner 2011). Nonetheless, excessive Na^+ is harmful for plants (Kronzucker et al. 2013) as was observed in the second round of experiments of the current study. Even though it was not a primary focus of the current project, however future experiments is recommended to first confirm the positive growth observed in this study and it would then assist in better understanding the gains and drawbacks of Na^+/K^+ relationship. Such knowledge would be useful for pea breeders to look into the opportunity of reducing costly K^+ based fertilizer with cheaper, energy efficient Na^+ based nutrients.

Next round of salt experiments revealed that *late3* and *late4* mutations show pronounced negative effect to salt stress compared to WT (Figure 7.12). In *A. thaliana*, Na^+ transporters such as NHXs and HKTs as well as ROS scavengers are known to play important role in salt stress resistance (Deinlein et al. 2014). The results from *A. thaliana* gene atlas showed

differential regulation of *AtCDK8* and *AtCYCC1* (Figure 7.1, Table A4.1). Besides, mediator complex component *AtMED25* was shown to exert salt resistance by acting downstream of DREB2A, ZFHD1 and MYB like transcription factors, and this role of *AtMED25* was stated as being conserved similar to that in an early embryophyte namely *Physomittrella patens* (Elfving et al. 2011). In a recent study, exposure of *A. thaliana* seedlings to salt stress resulted in upregulation of all the mediator complex genes (Dwivedi et al. 2017; Pasrija and Thakur 2012). The results of the present study are consistent with this as more severe salt stress phenotypes were observed in the *late3-2* and *late4-1* mutants. As mentioned earlier, *AtCDK8* was reported to act in ROS stress response in *A. thaliana* (Ng et al. 2013). Similarly, the predicted function of *MtCDK8* and *MtCYCC1* revealed their likely role in ROS metabolic process (Table 6.2, 6.3). Taken together these information, it can be speculated that *LATE3*/*PsCDK8* and *LATE4*/*PsCYCC1* act in a similar manner to render salt stress tolerance in pea, i.e., interacting with aforementioned orthologues of *A. thaliana* salt related TFs or being regulated by various abiotic stress related TF mentioned in chapter 6 (Table 6.4, 6.5).

7.5.6 Concluding remarks

Findings from this chapter have revealed understanding about the likely involvement of *LATE3*/*PsCDK8* and *LATE3*/*PsCYCC1* in broad regulatory processes needed to deal with various environmental factors which are eventually important for survival of a sessile organism, i.e., pea. Results have shown that both these genes are important for regulating general growth under both darkness and different light conditions. Besides, they seem to be important for mediating ambient temperature and light quality-temperature dependent flowering process. Likewise, these genes are likely to play role for carrying out plant defence response towards heat, UVB, mechanical injury and salt stress. The observed pattern of difference in response towards a certain treatment between *late3* and *late4* mutants could give hints towards existence of mediator independent *CDK8* function which is known in human (Firestein and Hahn 2009). Nevertheless, *AtCDK8* have already been reported as a potential integrator of numerous cellular signals which likely assist the plant to shift focus between development and stress response as necessary (Ng et al. 2013). Besides, information on involvement of *AtCYCC1* in biotic stress regulation is also available (Zhu et al. 2014). In general, the outcomes of this chapter involving *LATE3*/*PsCDK8* and *LATE3*/*PsCYCC1* are in line with the already available information and it guides towards understanding about the conserved role of

mediator complex across plant kingdom which probably have played very important role in evolutionary adaption of the primary food producers. Nevertheless, the results from the present chapter also laid down a strong platform for further research at a global transcriptome/proteome level in pea that could strengthen knowledge about specific role of *LATE3*/*PsCDK8* and *LATE3*/*PsCYCC1* for regulating response towards a certain environmental factor. As mediator complex act through RNA polymerase II which is known to regulate transcription of all protein and miRNA coding genes (Sikorski and Buratowski 2009), therefore *LATE3*/*PsCDK8* and *LATE3*/*PsCYCC1* could be important for rendering common response towards multiple environmental constraints simultaneously which can be unveiled by future investigations.

Chapter 8: General discussion

8.1 Key findings at a glance

In the current study, two novel late flowering mutants namely *late3* and *late4* have been functionally characterized. To this end, initial work involved inspection of various vegetative and reproductive traits that are affected in these mutants (Figure 3.1-3.11). It was found that mutation in the underlying genes affect various reproductive features such as transition from vegetative to reproductive phase, floral initiation, inflorescence and flower architecture, proper continuation of successive reproductive growth, floral fertility and timing of senescence. Besides, agronomically important traits such as pod size and shape, number of seeds/pod, number of seeds/plant, seed size and shape were also found be to be negatively affected by these mutations. Moreover, alteration in different vegetative traits such as leaflet size, petiole and proximal rachis length, stem thickness and length, internode length, aerial and basal branching, primary root length, overall shoot and root biomass were also observed. Impact of these mutations on such wide range of developmental features throughout various growth stages (Figure 8.1) gave strong hints that the underlying genes are global/universal regulators involved in various crucial regulatory processes in pea. Furthermore, the similarities in phenotypes between these mutants led to generation of hypothesis that *LATE3* and *LATE4* genes act in the same regulatory pathway.

Next, location of the *LATE3* and *LATE4* genes were determined by developing molecular markers and genetic maps using recombination data obtained from segregating F_2 mapping populations which were generated through cross between a pea cultivated line cv. Terese and *late3*, *late4* mutants (Figure 4.1-4.6). For this purpose, comparative genomics between pea and *M. truncatula* was successfully exploited which placed *LATE3* and *LATE4* loci within a narrow genetic interval co-segregating with other molecular markers in *PsLGIII* and *PsLGV* respectively. These regions were syntenic to *M. truncatula* chromosome 3 and 7 containing 62 and 54 annotated genes respectively (Table 5.1-5.2). In order to detect functional polymorphism between the mutants and wild-type in the relevant pea orthologous gene

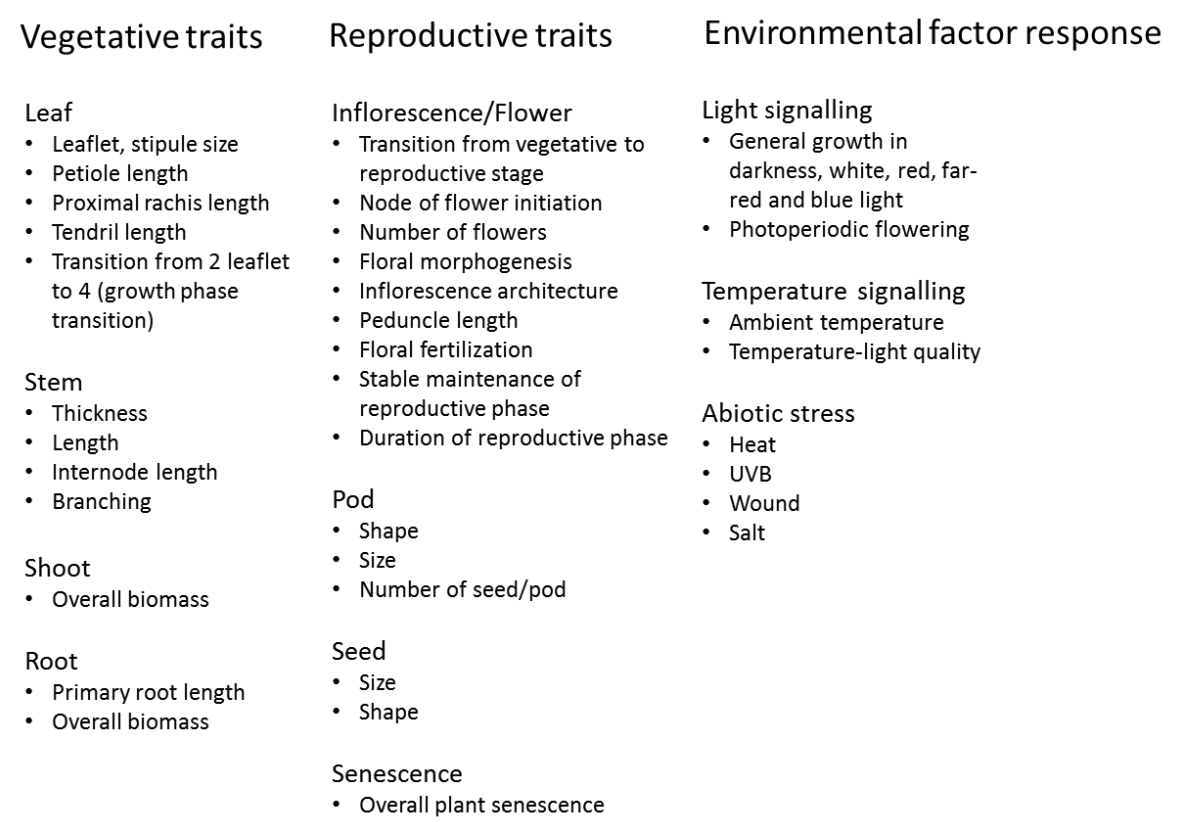


Figure 8.1. Role of *PsCDK8* and *PsCYCC1* genes in regulating various vegetative and reproductive traits and response towards different abiotic stress which have been captured either in a quantitative manner or observed during the present study.

within the defined interval, RNA sequencing was carried out which limited tedious task of sequencing many candidate genes directly, however the depth was found to be crucial. Based on RNA sequencing, phylogenetic analysis and direct sequencing of independent alleles of *late3* and *late4* mutations, it was found *LATE3* and *LATE4* genes are orthologues of *A. thaliana Cyclin Dependent Kinase 8* (*AtCDK8*) and *Cyclin C* (*AtCYCC1*) respectively. Both these genes are components of a kinase module called CDK8 which is part of the highly conserved large eukaryotic transcriptional regulator named mediator complex. Functional characterization of *cycc1* mutants has not been reported so far in any plant system and therefore the present study is the first of its kind.

SNPs resulted due to EMS mutagenesis in the *PsCDK8* and *PsCYCC1* genes were found to be associated with phenomenon such as alternative initiation of translation and transcription leading to generation of some common and very rare type of splice variants (Figure 5.6, 5.9).

Besides, these polymorphisms in *PsCDK8* and *PsCYCC1* were found to co-segregate with the mutant phenotype, thereby confirming initially determined genetic location of *LATE3* and *LATE4* genes (Figure 5.6 B, 5.8 B). In addition, analysis of gene atlas data in *A. thaliana*, *L. japonicus*, *M. truncatula*, *O. sativa* and *P. sativum* revealed more or less similar expression of these genes across various tissues (Figure A2.3-A2.4) providing hints about their previously hypothesized conserved global regulatory role. Besides, analysis of projected function of *MtCDK8* and *MtCYCC1* genes using the AraNET v2.0 database showed similarities with that of pea and also of *AtCYCC1* which was conducted in practice during this project (Table 6.2-6.3). Furthermore, studies on T-DNA insertion mutants of *AtCYCC1* gene unveiled similarities in phenotypic traits (Figure 5.11-5.12) that was already observed in pea which again suggested strongly about the conservation of the function of *CYCC1* gene in plant system. Thus, all these results formed the basis of identity of *LATE3* and *LATE4* as *PsCDK8* and *PsCYCC1*. As functionally significant mutations were detected in *PsCDK8* and *PsCYCC1* in multiple alleles of *late3* and *late4* mutants respectively similar to previously characterized pea mutants such as *late1* (Hecht et al. 2007), *gigas* (Hecht et al. 2011), *veg1* (Berbel et al. 2012), *sn* (Liew et al. 2014), *veg2* (Sussmilch et al. 2015) and *ppd* (Rubenach et al. 2017), therefore it gave strong evidence that they are causal genes for *LATE3* and *LATE4* loci.

Since PsCDK8 and PsCYCC1 are part of the same complex, i.e., CDK8 along with PsMED12 and PsMED13 and loss-of-function mutations of these genes are likely to have similar phenotypes, therefore one can argue why mutation occurred specifically in these genes. It is known that the sequences and methylation status surrounding the site of mutation strongly influences their likelihood of being alkylated by EMS (Henry et al. 2014) and this particular status were more congenial for *PsCDK8* and *PsCYCC1* than the other genes. Even though gDNA size of all of these genes are unknown, cDNA size of *PsCDK8* (~1.6 kb) and *PsCYCC1* (~1.1 Kb) are much smaller than *PsMED12a* (~7.3 kb), *PsMED12b* (~7.1 kb) and *PsMED13* (~6.2 kb) whereas the main functional domain of *PsCDK8* (kinase domain) and *PsCYCC1* (cyclin N domain) are larger than that of *PsMED12* (med12 domain, not shown) and *PsMED13* (med13 domain, not shown). Therefore, in comparison to the other three genes mutations in *PsCDK8*/*PsCYCC1* are more likely to affect their functions and generate significantly altered phenotype as observed in this study relative to WT (Chapter 3). Since both *PsCDK8* and *PsCYCC1* have a single copy in pea

(Chapter 5), therefore mutation in these genes are likely to be visible due to lack of additional copy of the same gene (exhibiting redundancy) which is the case specifically for *PsMED12*.

Given the scope and precedence, additional proof in validating *PsCDK8* and *PsCYCC1* as *LATE3* and *LATE4* respectively could be achieved in future through available reverse genetics methodologies in pea. For instance, the recently developed improved *Agrobacterium* mediated transformation system (Aftabi et al. 2017) could be exploited for generating gain of function (overexpression) and loss of function mutants (suppression) even though this system was not so successful for pea for long period of time due to low transformation rate (Schroeder et al. 1993; Švábová and Griga 2008). However, this system could not be used for the highly specific genome editing mechanism CRISPR/Cas9 technique which is yet to be generated for a species like pea (Song et al. 2016). In contrast, virus induced gene silencing has been used in pea limitedly and this particular approach may be taken (Grønlund et al. 2010). The TILLING mutant population in pea have been widely used for various studies, however this database do not have lines available for *PsCDK8* and *PsCYCC1* genes (Triques et al. 2007; Dalmais et al. 2008). Certain studies in the past mostly exploited complementation in *A. thaliana* using full length cDNA inserts of respective pea genes specifically when there was only a single mutant allele available e.g. *late2* (Ridge et al. 2016).

Candidate gene analysis also led to the generation of phylogenetic tree of SET domain proteins belonging to class I-V in *M. truncatula* and pea (Table 5.3). Due to their importance in epigenetic regulation of various reproductive and vegetative traits, the SET domain proteins are of key interest for scientists. Phylogenetic analysis suggested that the function of these genes is likely to be conserved between *A. thaliana* and the two legume species. Moreover, components of the *M. truncatula* and pea mediator complex were identified by protein blast analysis of *A. thaliana* mediator complex components and these results suggested higher degree of conservation of this important transcriptional regulator across various plant species (Table 5.4).

The observed similarities in many reproductive and vegetative traits (Figure 3.1-3.11, 8.1) between *late3* and *late4* mutants gave hints that they probably carry out the same regulatory process which was finally revealed by identifying *PsCDK8* and *PsCYCC1* as the causal genes respectively. In addition, genetic interaction experimentation showed complementarity

between *LATE3* and *LATE4* genes validating initial hypothesis of both of them acting in the same regulatory pathway (Figure 6.2). Knowledge on such functional relationship was further strengthened by observation of conserved strong physical interaction between *LATE3* and *LATE4* proteins through yeast two hybrid assay (Figure 6.3). In order to understand how *LATE3* and *LATE4* genes regulate flowering in pea, various other regulatory and genetic interaction studies were carried out. Results of regulatory interaction studies demonstrated that *LATE3* and *LATE4* genes act in pea flowering in a time dependent manner by promoting expression of known positive regulators of this pathway such as *FTa1*, *FTc* and various inflorescence and floral meristem identity genes such as *VEG1*, *UNI, PIM* and *VEG2* and also by repressing expression of negative regulators such as *LF* (Figure 6.6). These results suggest that *LATE3* and *LATE4* play positive role in higher cumulative expression of various floral activators such as *FTa1*, *FTc*, *PIM*, *UNI* and *VEG1* than the repression of floral repressors such as *LF* at 42 DAS which probably led to flowering in WT genotype. Besides, genetic interaction studies showed that *LATE3* and *LATE4* genes act upstream of negative regulators of pea flowering such as *LF* (*PsTFL1c*) and *DET* (*PsTFL1a*) (Figure 6.9-6.10). Similarly, it was also shown that *LATE3* and *LATE4* genes likely to act downstream of pea circadian clock gene *SN* (*PsLUX*) suggesting potential role of the gene in circadian clock mediated flowering in pea (Figure 6.13). In addition, both the genes were found to be crucial for photoperiod (Figure 6.4), ambient temperature and likely temperature-light quality (Figure 7.3) dependent flowering in pea.

Potential conservation of the function of CDK8 module in plant system have also been predicted in the present study as the various components of cell cycle and mediator complex were identified being the broader interaction network of the AtCDK8 and AtCYCC1 (Figure 6.14). By taking advantage of the large collection of plant transcription factors in the PlantTFDB v4.0 database, various TFs such those belonging mainly to AP2, BBR-BPC, bZIP, Dof, ERF, MYB families have been forecasted to regulate the function of the two mediator complex genes *CDK8* and *CYCC1* in pea and *M. truncatula* (Table 6.4-6.5). Thus, exploitation of various databases laid down groundwork for a time efficient, systems biology based approach in generation of hypothesis for future studies and also gaining further understanding about the results obtained from one plant system (e.g. pea) by cross-checking with another closely related plant system (e.g. *M. truncatula*).

Findings from the environmental factor experiments unravelled that *LATE3* and *LATE4* genes are essential for general growth under darkness and light, but not exclusively for photomorphogenesis process (Figure 7.2). Besides, both the genes are likely to regulate response towards UVB, heat, wound and salt stress (Figure 7.5, 7.7, 7.9, 7.11, 8.1). Therefore, it is highly likely that *PsCDK8* and *PsCYCC1* are universal transcription regulators which act not only to promote various vegetative and reproductive traits, but also participate actively in plant response towards various environmental factors – a feature that is absolutely essential for the survival of a sessile organism like pea. Thus, mediator complex components such as *CDK8* and *CYCC1* along with other subunits may have played significant role in evolutionary adaptation of pea under various conditions. Due to phylogenetically close relationship, similar role for *CDK8* and *CYCC1* genes can be expected in other legumes. Nevertheless, since pea and *M. truncatula* have shown certain species-specific variation for the origin of some of the subunits of mediator complex, therefore such kind of arrangement may have facilitated their adaptation to a certain condition.

RNA polymerase II carries out transcription of all the protein coding genes and miRNAs along with general transcription factors, TATA binding proteins and it associated factors and the mediator complex (Sikorski and Buratowski 2009; Dwivedi et al. 2017). Given such importance of mediator complex genes with regard to transcription of all protein coding genes among eukaryotes, various research have already revealed their global regulatory role in *A. thaliana* (Kidd et al. 2011; Samanta and Thakur 2015; Yang et al. 2016; Dwivedi et al. 2017; Malik et al. 2017). Specifically, *AtCDK8* was proposed to be a regulatory switch between growth and stress responses in plants (Ng et al. 2013). Similarly, *AtMED12* and *AtMED13* were defined as universal regulators of developmental phase transitions (Gillmor et al. 2014). Therefore, it can by hypothesized that phenotypes observed in *late3* and *late4* mutants could be direct effect of global regulatory role of *PsCDK8* and *PsCYCC1* affecting the transcription of underlying genes of various biological pathways. However, it could also be argued that various observed phenotypes of *late3*/*late4* mutants at different vegetative and reproductive stages could be a pleiotropic effect of late flowering/delayed transition of different stages of development. Normally, various assimilates are accumulated in flowers and seeds at the expense of vegetative organs during flowering and seed filling stage respectively (Burstin et al. 2007). Since, significantly delayed flowering and later on reduction in fertility/seed numbers were

observed in *late3*/*late4* mutants than WT, therefore it can be assumed that the source to sink resource allocation process has been significantly affected in the mutants. Various pleiotropic phenotypes have been reported previously in the loss-of-function *med12* mutants which showed delayed flowering (Imura et al. 2012).

It is known that individual CDK8 module components contribute differently to the cumulative effect of this kinase module. In animal system such as human, MED13 establishes a strong contact with the middle module of the core mediator followed by a weaker interaction of MED12 (Tsai et al. 2013). Unlike yeast, CDK8 and CYCC1 in human do not interact with the middle module directly. Therefore, MED13 and MED12 are considered the most important components for the role of CDK8 module. Deletion of MED13 completely eliminated interaction with core mediator whereas these interaction is weakened with deletion of *CDK8*. In *A. thaliana*, mutation in *MED13* and *MED12* was reported to have more effect on flowering initiation compared to *CDK8* (Zhu et al. 2014). The level of effect on the function of CDK8 module due to deletion of CYCC1 has not been shown in any studies. However, it can be assumed that deletion of CYCC1 is likely to have more effect compared to CDK8 since the former is the only mediator complex gene that has role in both transcriptional regulation as well as in cell cycle. During the present study, comparison of gene atlas in *M. truncatula* and pea revealed markedly higher expression of *CYCC1* gene across all tissues compared to *CDK8* which might be related to evolutionary divergence of the role of this gene that participate in two aforementioned important processes. Besides, higher level of response of the *late4* mutants were observed compared to *late3* mutants under certain abiotic stress treatments (chapter 7). These could be a result of dual impact on the CDK8 module function and progression of cell cycle occurring due to mutation in *LATE4/PsCYCC1* gene. Moreover, it was found in some cases (e.g. leaf area under various light conditions) that only *LATE3/PsCDK8* regulate the response, but not *LATE4/PsCYCC1*. This might happen due to mediator independent activity of CDK8 which is already known in human system (Firestein and Hahn 2009; Knuesel et al. 2009a; Knuesel et al. 2009b).

8.2 Proposed model of flowering with relevance to *LATE3* and *LATE4*

The major objective of the present study was to reveal the molecular identity of the *LATE3* and *LATE4* genes and understand how they might be regulating the process of flowering in

pea. As part of that, *LATE3* and *LATE4* have been identified as pea homologues of *CDK8* and *CYCC1* which are components of CDK8 module of the mediator complex (Chapter 3, 4 and 5). Besides, strong physical interaction and genetic complementarity between *LATE3* and *LATE4* genes revealed their functional dependence on each other. In addition, genetic and regulatory interaction studies showed that *LATE3* and *LATE4* broadly promote expression of already known key enhancers of flowering such as *FTa1* in leaf as well as *FTc*, *VEG1*, *PIM*, *UNI*, *VEG2* (time dependent) in apex (Chapter 6). In addition, both genetic and regulatory interaction studies revealed that *LATE3* and *LATE4* act negatively upon the floral repressor *LF*. Another genetic interaction study revealed a potential role for *LATE4* gene downstream of pea circadian clock gene *SN* where the latter is likely to negatively regulate the activity of the former. Even though, genetic interaction between *SN* and *LATE3* was not possible to conduct in the present study due to lack of segregation, but a similar type of interaction is highly possible due to similarities in the function of the two mediator complex genes. Besides, it is also hypothesized that other pea circadian clock genes *HR*, *PPD* and *DNE* would act in a same fashion upon *LATE3* and *LATE4* in order to mediate the clock driven flowering process (Figure 8.2).

It is known from the literature that core mediator made up of head, tail and middle module promote transcription of genes by assisting recruitment of RNA polymerase II whereas CDK8 module mainly act in transcriptional repression by blocking interaction between RNA polymerase II and core mediator (Elmlund et al. 2006; Yang et al. 2016), therefore these results suggest that *LATE3*/*PsCDK8* and *LATE4*/*PsCYCC1* are likely to promote the expression of aforementioned known positive regulators of flowering by repressing the function of yet known intermediate gene (s) (Figure 8.2). This regulation is most likely to be at the transcriptional level. On the other hand, *LATE3* and *LATE4* act negatively upon *LF* probably by regulating their transcription directly. However, it is also possible that they may repress the function of still unidentified positive regulator (s) of *LF* (Figure 8.2).

Knowledge about the role of mediator complex components is still at a very preliminary stage. However, studies in animal systems have unveiled role of CDK8 module components in promoting transcription rather than blocking this process (Donner et al. 2007; Donner et al. 2010; Belakavadi and Fondell 2010; Galbraith et al. 2010; Nemet et al. 2014). Therefore, it can be predicted that *LATE3* and *LATE4* also have similar role in pea. In such case, these genes are

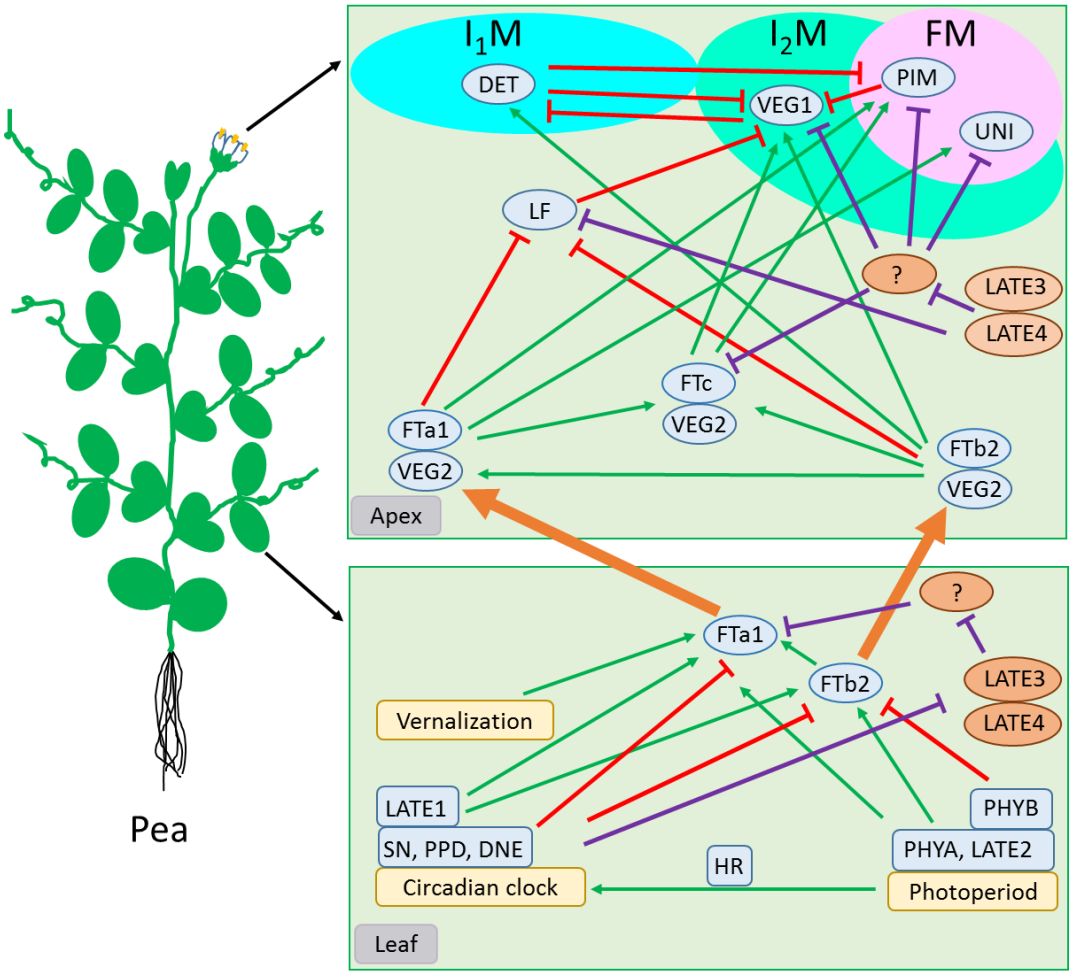


Figure 8.2. Hypothesized model of flowering time regulation in pea with relevance to *LATE3* and *LATE4* genes. Green arrows and red blunt headed lines represent known interactions revealed from previous models by Weller et al. 2009, Berbel et al. 2012 and Sussmilch et al., 2015. In contrast, purple blunt headed lines depict proposed interaction based on the results of present study which is also related to known role of CDK8 module components of highly conserved eukaryotic mediator complex. The question mark symbol represents unknown component (s) of the pathway, for simplicity only one such symbol has been shown both in leaf and apex which in reality could be multiple genes. Full names of the genes mentioned in text.

likely to positively regulate *FT*s and inflorescence/meristem identity genes by promoting their transcription. Similarly, these two CDK8 module genes would likely stimulate transcription of some intermediary gene that would in turn suppress function of *LF.*

Role of CDK8 module components in epigenetic regulation of gene expression have also been reported. It is known from the human system that MED12 interact with a histone methyltransferase in order to recruit H3K9 and H3K27 repressive epigenetic marks at target loci and thereby suppress their expression (Chaturvedi et al. 2009; Chaturvedi et al. 2012;

Ding et al. 2008). Such kind of epigenetic repression is more stable and stronger than that results from hindrance of association between RNA polymerase II and core mediator. Role of H3K9 and H3K27 in mediating seed-to-seedling as well as vegetative to reproductive transitions are also known (Gillmor et al. 2014). To this end, epigenetic regulation of key flowering genes such as *FT* and *FLC* in *A. thaliana* specially by the deposition of H3K27 methylation mark have already been discussed in chapter 1. Considering these aspects, *LATE3* and *LATE4* may affect the expression of the aforementioned key pea flowering genes through an epigenetic mechanism as well.

As the *late3* and *late4* mutants eventually flowered at a later time point and it was also found that various key flowering genes were induced (Figure 6.6), therefore it can be assumed that this was probably possible due to late activation of the canonical flowering pathway where photoperiod (Figure 6.4) plays role in a yet to be understood process in pea. The molecular basis of such flowering could be associated with the fact that PsMED12/13 (Table 5.4) may have still carried out the role of CDK8 module partially at a later time point even though activity of the entire complex was significantly affected by mutation in PsCDK8 and PsCYCC1. Besides, it was reported previously that CDK8 module genes *MED12* and *MED13* positively regulate juvenile to adult phase transition in *A. thaliana* by repressing expression of miR156 and promoting expression of miR172 and *SPL3/9* (Gillmor et al. 2014; Buendía-Monreal and Gillmor 2017). Therefore, expression of key age dependent pathway genes may have been similarly affected in *late3* and *late4* mutants at the onset of flowering and they ultimately flowered probably due to sufficient expression of pea homologues of miR172 and *SPL*s.

8.3 Potential future exploration on *LATE3* and *LATE4*

In the present study, it was observed that *late3* and *late4* mutants delay the formation of leaves having two leaflets to four leaflets which occur in the WT genotype at around 12 nodes (Wiltshire et al. 1994). This stage is being studied in a separate project using two other mutants and have been associated with juvenile vegetative to adult vegetative phase change (J. Vander Schoor, personal communication). In chapter 1, age dependent flowering pathway in *A. thaliana* has been briefly described where two micro-RNAs namely miR156 and miR172 play crucial role. As the age dependent phase transition and flowering is poorly understood

in pea and it seems that *LATE3* and *LATE4* are crucial in this regard, therefore it would be of interest to shed light on this aspect in future.

LATE3 and *LATE4* were found to be essential for successful maintenance of reproductive development. In *A. thaliana*, floral maintenance gene *LFY* was proposed to negatively regulate shoot meristem identity gene *TFL1* in order to prevent floral reversion (Parcy et al. 2002). In the present study, expression pattern of *UNI* (*PsLFY*) and *LF* (*PsTFL1c*) was in line with this information (chapter 6), however it would be of interest to study this process more carefully by using samples from each of the vegetative and reproductive stages in the mutants until senescence. Another important aspect which was noticed during the present study was the low rate of pod formation in *late3* and *late4* mutants even though they generated markedly higher numbers of flowers during their highly extended reproductive phase. This phenomenon has been hypothesized to be associated with failure in gametogenesis/fertilization/early embryogenesis process. To better understand how *LATE3* and *LATE4* act in the underlying regulatory processes which is directly related to yield outcome and economic profit, future research is recommended on this issue as well.

It was found in the current study that *late3* and *late4* mutants generate massive branches in comparison to the WT which indicated role of *LATE3* and *LATE4* in maintenance of apical dominance. Hormones like auxins, cytokinin, strigolactone as well as sugar status are known to play in this regard in *A. thaliana* (Barbier et al. 2015). However, this process is not understood well enough in pea. Determining the level of these hormones, sugar indicator trehalose 6-phosphate in the WT and mutants is going to reveal more knowledge in future on this very complex issue.

Once *LATE3* and *LATE4* were identified as *PsCDK8* and *PsCYCC1* which are components of the CDK8 module of the eukaryotic transcriptional regulator mediator complex, then the present study also revealed other members of the pea mediator complex. Therefore, knowledge on interaction between different members of the pea CDK8 module would be of interest. As the core mediator is known to promote transcription, therefore checking interaction of various components of the head, tail, middle and CDK8 module via yeast two hybrid/split luciferase complementation/GST pull down assay would unveil knowledge about the dynamic function of this crucial global transcriptional regulator in pea. Specially, interaction with already known

global regulators from studies on *A. thaliana* such as MED6, MED8, MED18 and MED25 could be undertaken as a priority. Such an intensive study was performed in yeast previously involving all the components of mediator complex (Guglielmi et al. 2004). In addition, genome wide association studies using gene bank accession of pea could assist in revealing the genetic basis of the role of *PsCDK8* and *PsCYCC1* as well as the entire mediator complex for evolutionary adaptation of this species in various conditions.

Comparative analysis of gene atlas data from the three different legumes, i.e., *L. japonicus*, *M. truncatula* and *P. sativum* has revealed similar expression pattern of the *CDK8* and *CYCC1* genes across various tissues and growth stages. Besides, the predicted functions of *MtCDK8* and *MtCYCC1* obtained from AraNET v2.0 showed similarities with various traits which were observed in the present study in pea and also for *AtCYCC1* in *A. thaliana*. As a proof-of-concept of the usefulness of such systems biology approach, future research could be designed by checking phenotypes in the transgenic line of both the genes in *M. truncatula* and then compare with the predicted information. A large collection of Tnt retrotransposon based insertion mutants in *M. truncatula* is already available which can be exploited in this regard (Tadege et al. 2008). Such studies would also unravel significant information about the role of both of these mediator complex genes in legumes.

It was found that *LATE3* and *LATE4* genes regulate response towards various environmental factors such as light, photoperiod, ambient temperature, heat, UVB, wound and salt. It would be of priority in future to conduct the abiotic stress experiments involving heat, UVB, wound and salinity to reveal the effect of such stress at cellular level through histochemical assay and microscopy which could not be performed due to limitation in time and scope during the current study. Moreover, emphasis should also be given to identify which of these environmental factors regulate the function of *LATE3*/*PsCDK8* and *LATE3*/*PsCYCC1* and this could be done by checking their expression upon a certain treatment. For this purpose, expression analysis of these genes using appropriate tissue samples needs to be carried out. Since, plants in field condition often endure multiple stress simultaneously, so it would be interesting to find out whether *LATE3* and *LATE4* are master regulators of abiotic stress response in pea by exposing them to various stresses at a time. Besides, the genetic basis of such responses and how *LATE3* and *LATE4* act in these pathways and any other pathways of interest would be better unveiled by global transcriptome via RNA sequencing or proteome

analysis through nanoLC-ESI-MS/MS. In addition, chip assay could be exploited in order to determine the broad scale genes with which PsCDK8 and PsCYCC1 interact and which of these pea homologs are known to be involved in pathways of interest in *A. thaliana*. Moreover, a combination of In vitro histone methyltransferase assay, chip assay and Q-RT-PCR could be used to determine the comparative enrichment level of repressive H3K9, H3K27 and activating H3K4, H3K36 methylation marks at the chromatins of key pea flowering genes *FTa1*, *FTc*, *UNI*, *VEG1*, *VEG2*, *UNI* and *PIM*. Results from such experiments using WT and mutant genotypes is likely to reveal whether *PsCDK8* and *PsCYCC1* is involved in epigenetic regulation of their target genes.

Another set of predictive studies identified putative TFs belonging to mainly AP2, bZIP, Dof, MYB, ERF etc. families that are likely to regulate function of the *MtCDK8* and *MtCYCC1* genes. Corresponding orthologues of these genes were also identified in pea. These TFs have role similar to various vegetative, reproductive and abiotic stress related traits that was observed for *PsCDK8* and *PsCYCC1* in the present study. Binding of these TFs at the identified motifs could be verified in future by Chip-PCR and promoter deletion studies. Besides, expression of *PsCDK8* and *PsCYCC1* genes is knockout and overexpressed lines of a particular TF gene could be carried out to show that the respective TF regulate function of these two mediator genes.

8.4 Concluding remarks

Results from this thesis significantly improved knowledge regarding the underlying molecular mechanism that drives flowering initiation in pea in relation to the two CDK8 module genes of the mediator complex *LATE3*/*PsCDK8* and *LATE4*/*PsCYCC1*. Besides, importance of these two genes for mediating various other reproductive and vegetative traits have been revealed. In addition, these genes were found to be important for regulating response to various developmental factors and abiotic stresses. The molecular basis of regulation of the various developmental and adaptive traits remains an area for future endeavour.

References

Abe M, Katsumata H, Komeda Y, Takahashi T (2003) Regulation of shoot epidermal cell differentiation by a pair of homeodomain proteins in Arabidopsis. Development 130 (4):635-643

Abe M, Kobayashi Y, Yamamoto S, Daimon Y, Yamaguchi A, Ikeda Y, Ichinoki H, Notaguchi M, Goto K, Araki T (2005) FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. Science 309 (5737):1052-1056. doi:DOI 10.1126/science.1115983

Abesingha G (2015) The effect of auxins on seed yield parameters in wheat, pea and canola grown under controlled environment and western canadian field conditions. MSc thesis, University of Alberta, Edmonton, AB

Acosta-Motos JR, Ortuno MF, Bernal-Vicente A, Diaz-Vivancos P, Sanchez-Blanco MJ, Hernandez JA (2017) Plant Responses to Salt Stress: Adaptive Mechanisms. Agronomy-Basel 7 (1). doi:ARTN 1810.3390/agronomy7010018

Adrian J, Torti S, Turck F (2009) From Decision to Commitment: The Molecular Memory of Flowering. Molecular Plant 2 (4):628-642. doi:https://doi.org/10.1093/mp/ssp031

Aftabi M, Negawo AT, Hassan F (2017) Improved Protocol for Agrobacterium-Mediated Transformation of Pea (Pisum sativum), vol 7. doi:10.4172/2168-9547.1000202

Agusti J, Lichtenberger R, Schwarz M, Nehlin L, Greb T (2011) Characterization of Transcriptome Remodeling during Cambium Formation Identifies MOL1 and RUL1 As Opposing Regulators of Secondary Growth. PLOS Genetics 7 (2). doi:ARTN e100131210.1371/journal.pgen.1001312

Alabadi D, Oyama T, Yanovsky MJ, Harmon FG, Mas P, Kay SA (2001) Reciprocal regulation between TOC1 and LHY/CCA1 within the Arabidopsis circadian clock. Science 293 (5531):880-883. doi:DOI 10.1126/science.1061320

Allen BL, Taatjes DJ (2015) The Mediator complex: a central integrator of transcription. Nature reviews Molecular cell biology 16 (3):155-166

Alqudah AM, Sharma R, Pasam RK, Graner A, Kilian B, Schnurbusch T (2014) Genetic Dissection of Photoperiod Response Based on GWAS of Pre-Anthesis Phase Duration in Spring Barley. PLOS One 9 (11):e113120. doi:10.1371/journal.pone.0113120

Alvarez-Venegas R, Pien S, Sadder M, Witmer X, Grossniklaus U, Avramova Z (2003) ATX-1, an Arabidopsis homolog of trithorax, activates flower homeotic genes. Current Biology 13 (8):627-637. doi:10.1016/S0960-9822(03)00243-4

Alves-Carvalho S, Aubert G, Carrère S, Cruaud C, Brochot A-L, Jacquin F, Klein A, Martin C, Boucherot K, Kreplak J, da Silva C, Moreau S, Gamas P, Wincker P, Gouzy J, Burstin J (2015) Full-length de novo assembly of RNA-seq data in pea (Pisum sativum L.) provides a gene expression atlas and gives insights into root nodulation in this species. The Plant Journal 84 (1):1-19. doi:10.1111/tpj.12967

Analysis of the genome sequence of the flowering plant Arabidopsis thaliana (2000). Nature 408(6814):796-815. doi:http://www.nature.com/nature/journal/v408/n6814/suppinfo/408796a0_S1.html

Anandhakumar J, Moustafa YW, Chowdhary S, Kainth AS, Gross DS (2016) Evidence for Multiple Mediator Complexes in Yeast Independently Recruited by Activated Heat Shock Factor. Molecular and Cellular Biology 36 (14):1943-1960. doi:10.1128/MCB.00005-16

Andersson CR, Helliwell CA, Bagnall DJ, Hughes TP, Finnegan EJ, Peacock WJ, Dennis ES (2008) The FLX gene of Arabidopsis is required for FRI-dependent activation of FLC expression. Plant and Cell Physiology 49 (2):191-200. doi:10.1093/pcp/pcm176

Andrau JC, van de Pasch L, Lijnzaad P, Bijma T, Koerkamp MG, van de Peppel J, Werner M, Holstege FC (2006) Genome-wide location of the coactivator mediator: Binding without activation and transient Cdk8 interaction on DNA. Molecular Cell 22 (2):179-192. doi:10.1016/j.molcel.2006.03.023

Andres F, Coupland G (2012) The genetic basis of flowering responses to seasonal cues. Nature reviews Genetics 13 (9):627-639. doi:10.1038/nrg3291

Angel A, Song J, Dean C, Howard M (2011) A Polycomb-based switch underlying quantitative epigenetic memory. Nature 476 (7358):105-109. doi:10.1038/nature10241

Arsovski AA, Galstyan A, Guseman JM, Nemhauser JL (2012) Photomorphogenesis. The Arabidopsis Book / American Society of Plant Biologists 10:e0147. doi:10.1199/tab.0147

Aubert G, Morin J, Jacquin F, Loridon K, Quillet MC, Petit A, Rameau C, Lejeune-Hénaut I, Huguet T, Burstin J (2006) Functional mapping in pea, as an aid to the candidate gene selection and for investigating synteny with the model legume Medicago truncatula. Theoretical and Applied Genetics 112 (6):1024-1041. doi:10.1007/s00122-005-0205-y

Aukerman MJ, Sakai H (2003) Regulation of flowering time and floral organ identity by a microRNA and its APETALA2-like target genes. The Plant Cell 15 (11):2730-2741. doi:DOI 10.1105/tpc.016238

Avery L, Wasserman S (1992) Ordering Gene-Function - the Interpretation of Epistasis in Regulatory Hierarchies. Trends in Genetics 8 (9):312-316. doi:Doi 10.1016/0168-9525(92)90140-Y

Bach G, Moskowitz S, Tieu P, Matynia A, Neufeld E (1993) Molecular analysis of Hurler syndrome in Druze and Muslim Arab patients in Israel: multiple allelic mutations of the IDUA gene in a small geographic area. American Journal of Human Genetics 53 (2):330

Backstrom S, Elfving N, Nilsson R, Wingsle G, Bjorklund S (2007) Purification of a plant mediator from Arabidopsis thaliana identifies PFT1 as the Med25 subunit. Molecular Cell 26 (5):717-729. doi:10.1016/j.molcel.2007.05.007

Balasubramanian S, Sureshkumar S, Lempe J, Weigel D (2006) Potent induction of Arabidopsis thaliana flowering by elevated growth temperature. PLOS Genetics 2 (7):e106. doi:10.1371/journal.pgen.0020106

Balderas-Hernández VE, Alvarado-Rodríguez M, Fraire-Velázquez S (2013) Conserved versatile master regulators in signalling pathways in response to stress in plants. AoB PLANTS 5:plt033-plt033. doi:10.1093/aobpla/plt033

Banaganapalli B, Mohammed K, Khan IA, Al‐Aama JY, Elango R, Shaik NA (2016) A Computational Protein Phenotype Prediction Approach to Analyze the Deleterious Mutations of Human MED12 Gene. Journal of Cellular Biochemistry 117 (9):2023-2035. doi:doi:10.1002/jcb.25499

Banyai G, Szilagyi Z, Baraznenok V, Khorosjutina O, Gustafsson CM, Lew DJ (2017) Cyclin C influences the timing of mitosis in fission yeast. Molecular Biology of the Cell 28 (13):1738-1744. doi:10.1091/mbc.e16-11-0787

Barber HN (1959) Physiological genetics of Pisum II. Heredity 13 (1):33-60

Barbier FF, Lunn JE, Beveridge CA (2015) Ready, steady, go! A sugar hit starts the race to shoot branching. Current Opinion in Plant Biology 25:39-45. doi:10.1016/j.pbi.2015.04.004

Basso K, Margolin AA, Stolovitzky G, Klein U, Dalla-Favera R, Califano A (2005) Reverse engineering of regulatory networks in human B cells. Nature Genetics 37 (4):382-390. doi:10.1038/ng1532

Bastianelli D, Grosjean F, Peyronnet C, Duparque M, Regnier JM (1998) Feeding value of pea (Pisum sativum, L.) - 1 Chemical composition of different categories of pea. Animal Science 67:609-619

Bastow R, Mylne JS, Lister C, Lippman Z, Martienssen RA, Dean C (2004) Vernalization requires epigenetic silencing of FLC by histone methylation. Nature 427 (6970):164-167. doi:10.1038/nature02269

Bates DO, Morris JC, Oltean S, Donaldson LF (2017) Pharmacology of Modulators of Alternative Splicing. Pharmacological Reviews 69 (1):63-79. doi:10.1124/pr.115.011239

Belakavadi M, Fondell JD (2010) Cyclin-Dependent Kinase 8 Positively Cooperates with Mediator To Promote Thyroid Hormone Receptor-Dependent Transcriptional Activation. Molecular and Cellular Biology 30 (10):2437-2448. doi:10.1128/Mcb.01541-09

Benavides FJ, Guénet J-L (2012) Chapter 1.4 - Mouse Genomics A2 - Hedrich, Hans J. In: The Laboratory Mouse (Second Edition). Academic Press, Boston, pp 57-90. doi:https://doi.org/10.1016/B978-0-12-382008-2.00004-0

Benito B, Haro R, Amtmann A, Cuin TA, Dreyer I (2014) The twins K+ and Na+ in plants. Journal of Plant Physiology 171 (9):723-731. doi:https://doi.org/10.1016/j.jplph.2013.10.014

Benlloch R, Berbel A, Ali L, Gohari G, Millán T, Madueño F (2015) Genetic control of inflorescence architecture in legumes. Frontiers in Plant Science 6:543. doi:10.3389/fpls.2015.00543

Benlloch R, Berbel A, Serrano-Mislata A, Madueno F (2007) Floral initiation and inflorescence architecture: a comparative view. Annals of Botany 100 (3):659-676. doi:10.1093/aob/mcm146

Bennett EJ, Roberts JA, Wagstaff C (2011) The role of the pod in seed development: strategies for manipulating yield. New Phytologist 190 (4):838-853. doi:10.1111/j.1469-8137.2011.03714.x

Berbel A, Ferrandiz C, Hecht V, Dalmais M, Lund OS, Sussmilch FC, Taylor SA, Bendahmane A, Ellis TH, Beltran JP, Weller JL, Madueno F (2012) VEGETATIVE1 is essential for development of the compound inflorescence in pea. Nature Communications 3:797. doi:10.1038/ncomms1801

Berger F, Hamamura Y, Ingouff M, Higashiyama T (2008) Double fertilization – caught in the act. Trends in Plant Science 13 (8):437-443. doi:https://doi.org/10.1016/j.tplants.2008.05.011

Bergey DR, Howe GA, Ryan CA (1996) Polypeptide signaling for plant defensive genes exhibits analogies to defense signaling in animals. Proceedings of the National Academy of Sciences of the United States of America 93 (22):12053-12058. doi:DOI 10.1073/pnas.93.22.12053

Berk AJ (1999) Activation of RNA polymerase II transcription. Current Opinion in Cell Biology11 (3):330-335. doi:Doi 10.1016/S0955-0674(99)80045-3

Bernal M, Verdaguer D, Badosa J, Abadía A, Llusià J, Peñuelas J, Núñez-Olivera E, Llorens L (2015) Effects of enhanced UV radiation and water availability on performance, biomass production and photoprotective mechanisms of Laurus nobilis seedlings. Environmental and Experimental Botany 109:264-275. doi:https://doi.org/10.1016/j.envexpbot.2014.06.016

Berr A, McCallum EJ, Ménard R, Meyer D, Fuchs J, Dong A, Shen WH (2010) Arabidopsis SET DOMAIN GROUP2 is required for H3K4 trimethylation and is crucial for both sporophyte and gametophyte development. The Plant Cell 22 (10):3232-3248. doi:10.1105/tpc.110.079962

Berr A, Shafiq S, Shen WH (2011) Histone modifications in transcriptional activation during plant development. BBA Gene Regulatory Mechanisms 1809 (10):567-576. doi:10.1016/j.bbagrm.2011.07.001

Beveridge CA, Batge SL, Ross JJ, Murfet IC (2001) Hormone physiology of pea mutants prevented from flowering by mutations gi or veg1. Physiologia Plantarum 113 (2):285-291. doi:DOI 10.1034/j.1399-3054.2001.1130217.x

Beveridge CA, Murfet IC (1996) The gigas mutant in pea is deficient in the floral stimulus. Physiologia Plantarum 96 (4):637-645

Bhattacharyya MK, Smith AM, Ellis THN, Hedley C, Martin C (1990) The wrinkled-seed character of pea described by Mendel is caused by a transposon-like insertion in a gene encoding starch-branching enzyme. Cell 60 (1):115-122. doi:https://doi.org/10.1016/0092-8674(90)90721-P

Bishopp A, Lehesranta S, Vatén A, Help H, El-Showk S, Scheres B, Helariutta K, Mähönen AP, Sakakibara H, Helariutta Y (2011) Phloem-transported cytokinin regulates polar auxin transport and maintains vascular pattern in the root meristem. Current Biology 21 (11):927-932

Bjorklund S, Antti H, Uddestrand I, Moritz T, Sundberg B (2007) Cross-talk between gibberellin and auxin in development of Populus wood: gibberellin stimulates polar auxin transport and has a common transcriptome with auxin. Plant Journal 52 (3):499-511. doi:10.1111/j.1365-313X.2007.03250.x

Black DL (1995) Finding splice sites within a wilderness of RNA. RNA 1 (8):763

Bleckmann A, Alter S, Dresselhaus T (2014) The beginning of a seed: regulatory mechanisms of double fertilization. Frontiers in Plant Science 5 (452). doi:10.3389/fpls.2014.00452

Blumwald E, Poole RJ (1985) Na+/H+ Antiport in Isolated Tonoplast Vesicles from Storage Tissue of Beta-Vulgaris. Plant Physiology 78 (1):163-167. doi:DOI 10.1104/pp.78.1.163

Bomblies K, Wang RL, Ambrose BA, Schmidt RJ, Meeley RB, Doebley J (2003) Duplicate FLORICAULA/LEAFY homologs zfl1 and zfl2 control inflorescence architecture and flower patterning in maize. Development 130 (11):2385-2395. doi:10.1242/dev.00457

Bonawitz ND, Kim JI, Tobimatsu Y, Ciesielski PN, Anderson NA, Ximenes E, Maeda J, Ralph J, Donohoe BS, Ladisch M, Chapple C (2014) Disruption of Mediator rescues the stunted growth of a lignin-deficient Arabidopsis mutant. Nature 509 (7500):376-380. doi:10.1038/nature13084

Booker J, Auldridge M, Wills S, McCarty D, Klee H, Leyser O (2004) MAX3/CCD7 is a carotenoid cleavage dioxygenase required for the synthesis of a novel plant signaling molecule. Current Biology 14 (14):1232-1238. doi:DOI 10.1016/j.cub.2004.06.061

Bordat A, Savois V, Nicolas M, Salse J, Chauveau A, Bourgeois M, Potier J, Houtin H, Rond C, Murat F, Marget P, Aubert G, Burstin J (2011) Translational Genomics in Legumes Allowed Placing In Silico 5460 Unigenes on the Pea Functional Map and Identified Candidate Genes in Pisum sativum L. G3: Genes, Genomes, Genetics 1 (2):93-103. doi:10.1534/g3.111.000349

Borggrefe T, Davis R, Erdjument-Bromage H, Tempst P, Kornberg RD (2002) A complex of the Srb8, -9, -10, and -11 transcriptional regulatory proteins from yeast. The Journal of Biological Chemistry 277 (46):44202-44207. doi:10.1074/jbc.M207195200

Borner R, Kampmann G, Chandler J, Gleissner R, Wisman E, Apel K, Melzer S (2000) A MADS domain gene involved in the transition to flowering in Arabidopsis. Plant Journal 24 (5):591-599. doi:DOI 10.1046/j.1365-313x.2000.00906.x

Bostock R, Stermer B (1989) Perspectives on wound healing in resistance to pathogens. Annual Review of Phytopathology 27 (1):343-371

Botstein D, Maurer R (1982) Genetic Approaches to the Analysis of Microbial Development. Annual Review of Genetics 16:61-83. doi:DOI 10.1146/annurev.ge.16.120182.000425

Bouquin T, Lasserre E, Pradier J, Pech JC, Balague C (1997) Wound and ethylene induction of the ACC oxidase melon gene CM-ACO1 occurs via two direct and independent transduction pathways. Plant Molecular Biology 35 (6):1029-1035. doi:Doi 10.1023/A:1005902226054

Bourbon H-M (2008) Comparative genomics supports a deep evolutionary origin for the large, four-module transcriptional mediator complex. Nucleic Acids Research 36 (12):3993-4008. doi:10.1093/nar/gkn349

Bouyer D, Roudier F, Heese M, Andersen ED, Gey D, Nowack MK, Goodrich J, Renou JP, Grini PE, Colot V, Schnittger A (2011) Polycomb Repressive Complex 2 Controls the Embryo-to-Seedling Phase Transition. PLOS Genetics 7 (3). doi:ARTN e100201410.1371/journal.pgen.1002014

Bowen ME, Henke K, Siegfried KR, Warman ML, Harris MP (2012) Efficient Mapping and Cloning of Mutations in Zebrafish by Low-Coverage Whole-Genome Sequencing. Genetics 190 (3):1017-U1229. doi:10.1534/genetics.111.136069

Bradley D, Ratcliffe O, Vincent C, Carpenter R, Coen E (1997) Inflorescence commitment and architecture in Arabidopsis. Science 275 (5296):80-83. doi:10.1126/science.275.5296.80

Bratzel F, López-Torrejón G, Koch M, Del Pozo JC, Calonje M (2010) Keeping cell identity in arabidopsis requires PRC1 RING-finger homologs that catalyze H2A monoubiquitination. Current Biology 20 (20):1853-1859. doi:10.1016/j.cub.2010.09.046

Braun N, de Saint Germain A, Pillot JP, Boutet-Mercey S, Dalmais M, Antoniadi I, Li X, Maia-Grondard A, Le Signor C, Bouteiller N, Luo D, Bendahmane A, Turnbull C, Rameau C (2012) The Pea TCP Transcription Factor PsBRC1 Acts Downstream of Strigolactones to Control Shoot Branching. Plant Physiology 158 (1):225-238. doi:10.1104/pp.111.182725

Brewer PB, Koltai H, Beveridge CA (2013) Diverse Roles of Strigolactones in Plant Development. Molecular Plant 6 (1):18-28. doi:10.1093/mp/sss130

Briggs WR, Olney MA (2001) Photoreceptors in plant photomorphogenesis to date. Five phytochromes, two cryptochromes, one phototropin, and one superchrome. Plant Physiology 125 (1):85-88. doi:DOI 10.1104/pp.125.1.85

Brogna S, Wen J (2009) Nonsense-mediated mRNA decay (NMD) mechanisms. Nature Structural &Amp; Molecular Biology 16:107. doi:10.1038/nsmb.1550

Bu Z, Yu Y, Li Z, Liu Y, Jiang W, Huang Y, Dong A-W (2014) Regulation of Arabidopsis Flowering by the Histone Mark Readers MRG1/2 via Interaction with CONSTANS to Modulate FT Expression. PLOS Genetics 10 (9):e1004617. doi:10.1371/journal.pgen.1004617

Bueckert RA, Wagenhoffer S, Hnatowich G, Warkentin TD (2015) Effect of heat and precipitation on pea yield and reproductive performance in the field. Canadian Journal of Plant Science 95 (4):629-639. doi:10.4141/Cjps-2014-342

Buendía-Monreal M, Gillmor CS (2017) Convergent repression of miR156 by sugar and the CDK8 module of Arabidopsis Mediator. Developmental Biology 423 (1):19-23. doi:https://doi.org/10.1016/j.ydbio.2017.01.007

Burset M, Seledtsov IA, Solovyev VV (2000) Analysis of canonical and non-canonical splice sites in mammalian genomes. Nucleic Acids Research 28 (21):4364-4375

Burstin J, Deniot G, Potier J, Weinachter C, Aubert G, Barranger A (2001) Microsatellite polymorphism in Pisum sativum. Plant Breeding 120 (4):311-317. doi:10.1046/j.1439-0523.2001.00608.x

Bussemaker HJ, Li H, Siggia ED (2001) Regulatory element detection using correlation with expression. Nature Genetics 27 (2):167-171. doi:Doi 10.1038/84792

Calvo SE, Pagliarini DJ, Mootha VK (2009) Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proceedings of the National Academy of Sciences 106 (18):7507-7512. doi:10.1073/pnas.0810916106

Campoli C, Drosse B, Searle I, Coupland G, von Korff M (2012) Functional characterisation of HvCO1, the barley (Hordeum vulgare) flowering time ortholog of CONSTANS. The Plant Journal : for Cell and Molecular Biology 69 (5):868-880. doi:10.1111/j.1365-313X.2011.04839.x

Campoli C, Pankin A, Drosse B, Casao CM, Davis SJ, Von Korff M (2013) HvLUX1 is a candidate gene underlying the early maturity 10 locus in barley: Phylogeny, diversity, and interactions with the circadian clock and photoperiodic pathways. New Phytologist 199 (4):1045-1059. doi:10.1111/nph.12346

Campoli C, Von Korff M (2014) Genetic Control of Reproductive Development in Temperate Cereals. Advances in Botanical Research, vol 72. doi:10.1016/B978-0-12-417162-6.00005-5

Cao Y, Dai Y, Cui S, Ma L (2008) Histone H2B monoubiquitination in the chromatin of Flowering Locus C regulates flowering time in Arabidopsis. The Plant Cell 20 (10):2586-2602. doi:10.1105/tpc.108.062760

Capovilla G, Schmid M, Pose D (2015) Control of flowering by ambient temperature. Journal of Experimental Botany 66 (1):59-69. doi:10.1093/jxb/eru416

Carrera I, Janody F, Leeds N, Duveau F, Treisman JE (2008) Pygopus activates Wingless target gene transcription through the mediator complex subunits Med12 and Med13. Proceedings of the National Academy of Sciences of the United States of America 105 (18):6644-6649. doi:10.1073/pnas.0709749105

Cerdan PD, Chory J (2003) Regulation of flowering time by light quality. Nature 423 (6942):881-885. doi:10.1038/nature01636

Cerdán PD, Chory J (2003) Regulation of flowering time by light quality. Nature 423 (6942):881

Chaturvedi CP, Hosey AM, Palii C, Perez-Iratxeta C, Nakatani Y, Ranish JA, Dilworth FJ, Brand M (2009) Dual role for the methyltransferase G9a in the maintenance of beta-globin gene transcription in adult erythroid cells. Proceedings of the National Academy of Sciences of the United States of America 106 (43):18303-18308. doi:10.1073/pnas.0906769106

Chaturvedi CP, Somasundaram B, Singh K, Carpenedo RL, Stanford WL, Dilworth FJ, Brand M (2012) Maintenance of gene silencing by the coordinate action of the H3K9 methyltransferase G9a/KMT1C and the H3K4 demethylase Jarid1a/KDM5A. Proceedings of the National Academy of Sciences of the United States of America 109 (46):18845-18850. doi:10.1073/pnas.1213951109

Chen R, Jiang HL, Li L, Zhai QZ, Qi LL, Zhou WK, Liu XQ, Li HM, Zheng WG, Sun JQ, Li CY (2012) The Arabidopsis Mediator Subunit MED25 Differentially Regulates Jasmonate and Abscisic Acid Signaling through Interacting with the MYC2 and ABI5 Transcription Factors. The Plant Cell 24 (7):2898-2916. doi:10.1105/tpc.112.098277

Cheng JZ, Zhou YP, Lv TX, Xie CP, Tian CE (2017) Research progress on the autonomous flowering time pathway in Arabidopsis. Physiology and Molecular Biology of Plants23 (3):477-485. doi:10.1007/s12298-017-0458-3

Choi K, Park C, Lee J, Oh M, Noh B, Lee I (2007) Arabidopsis homologs of components of the SWR1 complex regulate flowering and plant development. Development 134 (10):1931-1941. doi:10.1242/dev.001891

Collard BCY, Jahufer MZZ, Brouwer JB, Pang ECK (2005) An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. Euphytica 142 (1-2):169-196. doi:10.1007/s10681-005-1681-5

Conaway RC, Sato S, Tomomori-Sato C, Yao TT, Conaway JW (2005) The mammalian Mediator complex and its role in transcriptional regulation. Trends in Biochemical Sciences 30 (5):250-255. doi:10.1016/j.tibs.2005.03.002

Coneva V, Zhu T, Colasanti J (2007) Expression differences between normal and indeterminate1 maize suggest downstream targets of ID1, a floral transition regulator in maize. Journal of Experimental Botany 58 (13):3679-3693. doi:10.1093/jxb/erm217

Corbesier L, Vincent C, Jang SH, Fornara F, Fan QZ, Searle I, Giakountis A, Farrona S, Gissot L, Turnbull C, Coupland G (2007) FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. Science 316 (5827):1030-1033. doi:10.1126/science.1141752

Cornic G, Ghashghaie J (1991) Effect of Temperature on Net Co2 Assimilation and Photosystem-Ii Quantum Yield of Electron-Transfer of French Bean (Phaseolus-Vulgaris L) Leaves during Drought Stress. Planta 185 (2):255-260

Côrte-Real HBSM, Dixon DR, Holland PWH (1994) Intron-targeted PCR: a new approach to survey neutral DNA polymorphism in bivalve populations. Marine Biology 120 (3):407-413. doi:10.1007/BF00680214

Crawford NM (1995) Nitrate - Nutrient and Signal for Plant-Growth. The Plant Cell 7 (7):859-868. doi:DOI 10.1105/tpc.7.7.859

Creelman RA, Mullet JE (1997) Biosynthesis and action of jasmonates in plants. Annual Review of Plant Physiology and Plant Molecular Biology 48:355-381. doi:DOI 10.1146/annurev.arplant.48.1.355

Crevillen P, Dean C (2011) Regulation of the floral repressor gene FLC: the complexity of transcription in a chromatin context. Current Opinion in Plant Biology 14 (1):38-44. doi:DOI 10.1016/j.pbi.2010.08.015

Cronk Q, Ojeda I, Pennington RT (2006) Legume comparative genomics: progress in phylogenetics and phylogenomics. Current Opinion in Plant Biology 9 (2):99-103. doi:10.1016/j.pbi.2006.01.011

Dalmais M, Schmidt J, Le Signor C, Moussy F, Burstin J, Savois V, Aubert G, Brunaud V, de Oliveira Y, Guichard C, Thompson R, Bendahmane A (2008) UTILLdb, a Pisum sativum in silico forward and reverse genetics tool. Genome Biology 9 (2):R43-R43. doi:10.1186/gb-2008-9-2-r43

Daviere JM, Achard P (2013) Gibberellin signaling in plants. Development 140 (6):1147-1151. doi:DOI 10.1242/dev.087650

Davletova S, Schlauch K, Coutu J, Mittler R (2005) The zinc-finger protein Zat12 plays a central role in reactive oxygen and abiotic stress signaling in Arabidopsis. Plant Physiology 139 (2):847-856. doi:10.1104/pp.105.068254

De Angioletti M, Lacerra G, Sabato V, Carestia C (2004) β +45 G → C: a novel silent β -thalassaemia mutation, the first in the Kozak sequence. British Journal of Haematology 124 (2):224-231. doi:10.1046/j.1365-2141.2003.04754.x

De Conti L, Baralle M, Buratti E (2013) Exon and intron definition in pre-mRNA splicing. Wiley Interdisciplinary Reviews: RNA 4 (1):49-60. doi:10.1002/wrna.1140

de Folter S, Immink RG (2011) Yeast protein-protein interaction assays and screens. Methods in Molecular Biology 754:145-165. doi:10.1007/978-1-61779-154-3_8

De Lucia F, Crevillen P, Jones AM, Greb T, Dean C (2008) A PHD-polycomb repressive complex 2 triggers the epigenetic silencing of FLC during vernalization. Proceedings of the National Academy of Sciences of the United States of America 105 (44):16831-16836. doi:10.1073/pnas.0808687105

Deal RB, Topp CN, McKinney EC, Meagher RB (2007) Repression of flowering in Arabidopsis requires activation of FLOWERING LOCUS C expression by the histone variant H2A.Z. The Plant Cell 19 (1):74-83. doi:10.1105/tpc.106.048447

Deinlein U, Stephan AB, Horie T, Luo W, Xu G, Schroeder JI (2014) Plant salt-tolerance mechanisms. Trends in Plant Science 19 (6):371-379. doi:10.1016/j.tplants.2014.02.001

Demason DA, Chawla R (2006) Auxin/gibberellin interactions in pea leaf morphogenesis. Botanical Journal of the Linnean Society 150 (1):45-59. doi:DOI 10.1111/j.1095-8339.2006.00491.x

DeMason DA, Villani PJ (2001) Genetic control of leaf development in pea (Pisum sativum). International Journal of Plant Sciences 162 (3):493-511. doi:Doi 10.1086/320137

Dennis ES, Peacock WJ (2007) Epigenetic regulation of flowering. Current Opinion in Plant Biology 10 (5):520-527. doi:http://dx.doi.org/10.1016/j.pbi.2007.06.009

Deulvot C, Charrel H, Marty A, Jacquin F, Donnadieu C, Lejeune-Hénaut I, Burstin J, Aubert G (2010) Highly-multiplexed SNP genotyping for genetic mapping and germplasm diversity studies in pea. BMC Genomics 11 (1):468. doi:10.1186/1471-2164-11-468

Dietz HC, Valle D, Francomano CA, Kendzior RJ, Pyeritz RE, Cutting GR (1993) The skipping of constitutive exons in vivo induced by nonsense mutations. Science 259 (5095):680-683

Ding N, Zhou H, Esteve PO, Chin HG, Kim S, Xu X, Joseph SM, Friez MJ, Schwartz CE, Pradhan S, Boyer TG (2008) Mediator links epigenetic silencing of neuronal gene expression with x-linked mental retardation. Molecular Cell 31 (3):347-359. doi:10.1016/i.molcel.2008.05.023

Ding XD, Richter T, Chen M, Fujii H, Seo YS, Xie MT, Zheng XW, Kanrar S, Stevenson RA, Dardick C, Li Y, Jiang H, Zhang Y, Yu FH, Bartley LE, Chern M, Bart R, Chen XH, Zhu LH, Farmerie WG, Gribskov M, Zhu JK, Fromm ME, Ronald PC, Song WY (2009) A Rice Kinase-Protein Interaction Map. Plant Physiology 149 (3):1478-1492. doi:10.1104/pp.108.128298

Dinneny JR, Yanofsky MF (2005) Drawing lines and borders: how the dehiscent fruit of Arabidopsis is patterned. BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology 27 (1):42-49

Distelfeld A, Li C, Dubcovsky J (2009) Regulation of flowering in temperate cereals. Current Opinion in Plant Biology 12 (2):178-184. doi:10.1016/j.pbi.2008.12.010

Doi K, Izawa T, Fuse T, Yamanouchi U, Kubo T, Shimatani Z, Yano M, Yoshimura A (2004) Ehd1, a B-type response regulator in rice, confers short-day promotion of flowering and controls FT-like gene expression independently of Hd1. Genes & Development 18 (8):926-936. doi:10.1101/gad.1189604

Doležel J, Vrána J, Šafář J, Bartoš J, Kubaláková M, Šimková H (2012) Chromosomes in the flow to simplify genome analysis. Functional & Integrative Genomics 12 (3):397-416. doi:10.1007/s10142-012-0293-0

Domagalska MA, Schomburg FM, Amasino RM, Vierstra RD, Nagy F, Davis SJ (2007) Attenuation of brassinosteroid signaling enhances FLC expression and delays flowering. Development 134 (15):2841-2850. doi:10.1242/dev.02866

Dong Z, Danilevskaya O, Abadie T, Messina C, Coles N, Cooper M (2012) A gene regulatory network model for Floral transition of the shoot apex in maize and its dynamic modeling. PLOS One 7 (8). doi:10.1371/journal.pone.0043450

Donner AJ, Ebmeier CC, Taatjes DJ, Espinosa JM (2010) CDK8 is a positive regulator of transcriptional elongation within the serum response network. Nature Structural & Molecular Biology 17 (2):194-U199. doi:10.1038/nsmb.1752

Donner AJ, Szostek S, Hoover JM, Espinosa JM (2007) CDK8 is a stimulus-specific positive coregulator of p53 target genes. Molecular Cell 27 (1):121-133. doi:10.1016/j.molcel.2007.05.026

Dortay H, Gruhn N, Pfeifer A, Schwerdtner M, Schmulling T, Heyl A (2008) Toward an interaction map of the two-component signaling pathway of Arabidopsis thaliana. Journal of Proteome Research 7 (9):3649-3660. doi:10.1021/pr0703831

Dresselhaus T, Sprunck S, Wessel GM (2016) Fertilization Mechanisms in Flowering Plants. Current Biology : CB 26 (3):R125-R139. doi:10.1016/j.cub.2015.12.032

Du HW, Zhu JX, Su H, Huang M, Wang HW, Ding SC, Zhang BL, Luo A, Wei SD, Tian XH, Xu YB (2017) Bulked Segregant RNA-seq Reveals Differential Expression and SNPs of Candidate Genes Associated with Waterlogging Tolerance in Maize. Frontiers in Plant Science 8. doi:ARTN 102210.3389/fpls.2017.01022

Duan L, Sebastian J, Dinneny JR (2015) Salt-Stress Regulation of Root System Growth and Architecture in Arabidopsis Seedlings. In: Estevez JM (ed) Plant Cell Expansion: Methods and Protocols. Springer New York, New York, NY, pp 105-122. doi:10.1007/978-1-4939-1902-4_10

Duarte J, Rivière N, Baranger A, Aubert G, Burstin J, Cornet L, Lavaud C, Lejeune-Hénaut I, Martinant J-P, Pichon J-P, Pilet-Nayel M-L, Boutet G (2014) Transcriptome sequencing for high throughput SNP development and genetic mapping in Pea. BMC Genomics 15:126-126. doi:10.1186/1471-2164-15-126

Dunford RP, Griffiths S, Christodoulou V, Laurie DA (2005) Characterisation of a barley (Hordeum vulgare L.) homologue of the Arabidopsis flowering time regulator GIGANTEA. TAG Theoretical and applied genetics Theoretische und angewandte Genetik 110 (5):925-931. doi:10.1007/s00122-004-1912-5

Dwivedi N, Kumar V, Thakur JK (2017) Convergence of Stress ‐ Induced Hormone Signaling Pathways on a Transcriptional Co ‐ Factor. Mechanism of Plant Hormone Signaling under Stress:285-317

Eddy SR (2004) Where did the BLOSUM62 alignment score matrix come from? Nature Biotechnology 22:1035. doi:10.1038/nbt0804-1035

https://www.nature.com/articles/nbt0804-1035#supplementary-information

Edwards EJ, Chatelet DS, Sack L, Donoghue MJ (2014) Leaf life span and the leaf economic spectrum in the context of whole plant architecture. Journal of Ecology 102 (2):328-336. doi:10.1111/1365-2745.12209

Elfving N, Davoine C, Benlloch R, Blomberg J, Brännström K, Müller D, Nilsson A, Ulfstedt M, Ronne H, Wingsle G, Nilsson O, Björklund S (2011) The Arabidopsis thaliana Med25 mediator subunit integrates environmental cues to control plant development. Proceedings of the National Academy of Sciences of the United States of America 108 (20):8245-8250. doi:10.1073/pnas.1002981108

Ellis T (2009) Legume genetics and breeding: the example of pea. Grain Legumes 51:14-15

Elmlund H, Baraznenok V, Lindahl M, Samuelsen CO, Koeck PJB, Holmberg S, Hebert H, Gustafsson CM (2006) The cyclin-dependent kinase 8 module sterically blocks Mediator interactions with RNA polymerase II. Proceedings of the National Academy of Sciences 103 (43):15788-15793. doi:10.1073/pnas.0607483103

Exner V, Alexandre C, Rosenfeldt G, Alfarano P, Nater M, Caflisch A, Gruissem W, Batschauer A, Hennig L (2010) A Gain-of-Function Mutation of Arabidopsis CRYPTOCHROME1 Promotes Flowering. Plant Physiology 154 (4):1633-1645. doi:10.1104/pp.110.160895

Farrona S, Thorpe FL, Engelhorn J, Adrian J, Dong X, Sarid-Krebs L, Goodrich J, Turcka F (2011) Tissue-specific expression of flowering locus t in arabidopsis is maintained independently of polycomb group protein repression. The Plant Cell 23 (9):3204-3214. doi:10.1105/tpc.111.087809

Faure S, Turner AS, Gruszka D, Christodoulou V, Davis SJ, von Korff M, Laurie DA (2012) Mutation at the circadian clock gene EARLY MATURITY 8 adapts domesticated barley (Hordeum vulgare) to short growing seasons. Proceedings of the National Academy of Sciences 109 (21):8328-8333. doi:10.1073/pnas.1120496109

Feng P, Guo H, Chi W, Chai X, Sun X, Xu X, Ma J, Rochaix J-D, Leister D, Wang H, Lu C, Zhang L (2016) Chloroplast retrograde signal regulates flowering. Proceedings of the National Academy of Sciences 113 (38):10708-10713. doi:10.1073/pnas.1521599113

Ferrandiz C, Gu Q, Martienssen R, Yanofsky MF (2000) Redundant regulation of meristem identity and plant architecture by FRUITFULL, APETALA1 and CAULIFLOWER. Development 127 (4):725-734

Fichtner F, Barbier FF, Feil R, Watanabe M, Annunziata MG, Chabikwa TG, Höfgen R, Stitt M, Beveridge CA, Lunn JE (2017) Trehalose 6-phosphate is involved in triggering axillary bud outgrowth in garden pea (Pisum sativum L.). The Plant Journal 92 (4):611-623. doi:10.1111/tpj.13705

Finnegan EJ, Dennis ES (2007) Vernalization-Induced Trimethylation of Histone H3 Lysine 27 at FLC Is Not Maintained in Mitotically Quiescent Cells. Current Biology 17 (22):1978-1983. doi:10.1016/j.cub.2007.10.026

Firestein R, Hahn WC (2009) Revving the Throttle on an Oncogene: CDK8 Takes the Driver Seat. Cancer Research 69 (20):7899-7901. doi:10.1158/0008-5472.Can-09-1704

Flanagan PM, Kelleher RJ, 3rd, Sayre MH, Tschochner H, Kornberg RD (1991) A mediator required for activation of RNA polymerase II transcription in vitro. Nature 350 (6317):436-438. doi:10.1038/350436a0

Fleet CM, Sun TP (2005) A DELLAcate balance: the role of gibberellin in plant morphogenesis. Current Opinion in Plant Biology 8 (1):77-85. doi:10.1016/j.pbi.2004.11.015

Ford R, Le Roux K, Itman C, Brouwer JB, Taylor PWJ (2002) Diversity analysis and genotyping in Pisum with sequence tagged microsatellite site (STMS) primers. Euphytica 124 (3):397-405. doi:10.1023/a:1015752907108

Forde BG (2002) The role of long-distance signalling in plant responses to nitrate and other nutrients. Journal of Experimental Botany 53 (366):39-43. doi:DOI 10.1093/jexbot/53.366.39

Fornara F, de Montaigu A, Coupland G (2010) SnapShot: control of flowering in Arabidopsis. Cell 141 (3):550-550. e552

Foucher F, Morin J, Courtiade J, Cadioux S, Ellis N, Banfield MJ, Rameau C (2003) DETERMINATE and LATE FLOWERING are two TERMINAL FLOWER1/CENTRORADIALIS homologs that control two distinct phases of flowering initiation and development in pea. The Plant Cell 15 (11):2742-2754. doi:10.1105/tpc.015701

Frances S, White MJ, Edgerton MD, Jones AM, Elliott RC, Thompson WF (1992) Initial Characterization of a Pea Mutant with Light-Independent Photomorphogenesis. The Plant Cell 4 (12):1519-1530

Fraser HB, Hirsh AE, Wall DP, Eisen MB (2004) Coevolution of gene expression among interacting proteins. Proceedings of the National Academy of Sciences of the United States of America 101 (24):9033-9038. doi:10.1073/pnas.0402591101

Fu R, Zhang M, Zhao Y, He X, Ding C, Wang S, Feng Y, Song X, Li P, Wang B (2017) Identification of Salt Tolerance-related microRNAs and Their Targets in Maize (Zea mays L.) Using High-throughput Sequencing and Degradome Analysis. Frontiers in Plant Science 8 (864). doi:10.3389/fpls.2017.00864

Furness N, Jolliffe P, Upadhyaya M (2005) Competitive interactions in mixtures of broccoli and Chenopodium album grown at two UV‐B radiation levels under glasshouse conditions. Weed Research 45 (6):449-459

Galbraith MD, Donner AJ, Espinosa JM (2010) CDK8: A positive regulator of transcription. Transcription 1 (1):4-12. doi:10.4161/trns.1.1.12373

Galvao VC, Horrer D, Kuttner F, Schmid M (2012) Spatial control of flowering by DELLA proteins in Arabidopsis thaliana. Development 139 (21):4072-4082. doi:DOI 10.1242/dev.080879

Garg R, Jain M (2013) RNA-Seq for Transcriptome Analysis in Non-model Plants. In: Rose RJ (ed) Legume Genomics: Methods and Protocols. Humana Press, Totowa, NJ, pp 43-58. doi:10.1007/978-1-62703-613-9_4

Ge H, Liu ZH, Church GM, Vidal M (2001) Correlation between transcriptome and interactome mapping data from Saccharomyces cerevisiae. Nature Genetics 29 (4):482-486. doi:DOI 10.1038/ng776

Gibson RA, Hajianpour A, Murer-Orlando M, Buchwald M, Mathew CG (1993) A nonsense mutation and exon skipping in the Fanconi anaemia group C gene. Human Molecular Genetics2 (6):797-799

Gillmor CS, Park MY, Smith MR, Pepitone R, Kerstetter RA, Poethig RS (2010) The MED12-MED13 module of Mediator regulates the timing of embryo patterning in Arabidopsis. Development (Cambridge, England) 137 (1):113-122. doi:10.1242/dev.043174

Gillmor CS, Silva-Ortega CO, Willmann MR, Buendía-Monreal M, Poethig RS (2014) The Arabidopsis Mediator CDK8 module genes CCT (MED12) and GCT (MED13) are global regulators of developmental phase transitions. Development 141 (23):4580-4589

Giri, Anju, "Effect of acute heat stress on nutrient uptake by plant roots" (2013).Theses and Dissertations. 82. http://utdr.utoledo.edu/theses-dissertations/82

Gnan S, Priest A, Kover PX (2014) The Genetic Basis of Natural Variation in Seed Size and Seed Number and Their Trade-Off Using Arabidopsis thaliana MAGIC Lines. Genetics 198 (4):1751-1758. doi:10.1534/genetics.114.170746

Golembeski GS, Imaizumi T (2015) Photoperiodic Regulation of Florigen Function in Arabidopsis thaliana. The Arabidopsis Book:e0178. doi:10.1199/tab.0178

Gong Y-m, Xu S-c, Mao W-h, Hu Q-z, Zhang G-w, Ding J, Li Y-d (2010) Developing new SSR markers from ESTs of pea (Pisum sativum L.). Journal of Zhejiang University Science B 11 (9):702-707. doi:10.1631/jzus.B1000004

Gonzalez N, Vanhaeren H, Inzé D (2012) Leaf size control: complex coordination of cell division and expansion. Trends in Plant Science 17 (6):332-340. doi:https://doi.org/10.1016/j.tplants.2012.02.003

Griffiths S, Dunford RP, Coupland G, Laurie DA (2003) The evolution of CONSTANS-like gene families in barley, rice, and Arabidopsis. Plant Physiology 131 (4):1855-1867. doi:10.1104/pp.102.016188

Grimstad SO (1993) The Effect of a Daily Low-Temperature Pulse on Growth and Development of Greenhouse Cucumber and Tomato Plants during Propagation. Sci Hortic-Amsterdam 53 (1-2):53-62. doi:Doi 10.1016/0304-4238(93)90137-F

Gu W, Malik S, Ito M, Yuan CX, Fondell JD, Zhang XL, Martinez E, Qin J, Roeder RG (1999) A novel human SRB/MED-containing cofactor complex, SMCC, involved in transcription regulation. Molecular Cell 3 (1):97-108. doi:Doi 10.1016/S1097-2765(00)80178-1

Gu X, Jiang D, Wang Y, Bachmair A, He Y (2009) Repression of the floral transition via histone H2B monoubiquitination. Plant Journal 57 (3):522-533. doi:10.1111/j.1365-313X.2008.03709.x

Gu X, Wang Y, He Y (2013) Photoperiodic Regulation of Flowering Time through Periodic Histone Deacetylation of the Florigen Gene FT. PLOS Biology 11 (9):e1001649. doi:10.1371/journal.pbio.1001649

Guglielmi B, van Berkum NL, Klapholz B, Bijma T, Boube M, Boschiero C, Bourbon H-M, Holstege FCP, Werner M (2004) A high resolution protein interaction map of the yeast Mediator complex. Nucleic Acids Research 32 (18):5379-5391. doi:10.1093/nar/gkh878

Guo A-Y, Chen X, Gao G, Zhang H, Zhu Q-H, Liu X-C, Zhong Y-F, Gu X, He K, Luo J (2008a) PlantTFDB: a comprehensive plant transcription factor database. Nucleic Acids Research 36 (Database issue):D966-D969. doi:10.1093/nar/gkm841

Guo L, Yu Y, Law JA, Zhang X (2010) Set domain group2 is the major histone H3 lysine 4 trimethyltransferase in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America 107 (43):18557-18562. doi:10.1073/pnas.1010478107

Guo M, Thomas J, Collins G, Timmermans MC (2008b) Direct repression of KNOX loci by the ASYMMETRIC LEAVES1 complex of Arabidopsis. The Plant Cell 20 (1):48-58. doi:10.1105/tpc.107.056127

Hahn A, Bublak D, Schleiff E, Scharf K-D (2011) Crosstalk between Hsp90 and Hsp70 chaperones and heat stress transcription factors in tomato. The Plant Cell 23 (2):741-755

Hahn A, Kilian J, Mohrholz A, Ladwig F, Peschke F, Dautel R, Harter K, Berendzen KW, Wanke D (2013) Plant core environmental stress response genes are systemically coordinated during abiotic stresses. International Journal of Molecular Sciences 14 (4):7617-7641. doi:10.3390/ijms14047617

Hamamura Y, Nagahara S, Higashiyama T (2012) Double fertilization on the move. Current Opinion in Plant Biology 15 (1):70-77. doi:10.1016/j.pbi.2011.11.001

Han Y-J, Song P-S, Kim J-I (2007) Phytochrome-mediated photomorphogenesis in plants. Journal of Plant Biology 50 (3):230-240. doi:10.1007/bf03030650

Hanano S, Goto K (2011) Arabidopsis terminal flower1 is involved in the regulation of flowering time and inflorescence development through transcriptional repression. The Plant Cell 23 (9):3172-3184. doi:10.1105/tpc.111.088641

Hasanuzzaman M, Nahar K, Alam MM, Roychowdhury R, Fujita M (2013) Physiological, biochemical, and molecular mechanisms of heat stress tolerance in plants. International Journal of Molecular Sciences 14 (5):9643-9684. doi:10.3390/ijms14059643

Hayama R, Yokoi S, Tamaki S, Yano M, Shimamoto K (2003) Adaptation of photoperiodic control pathways produces short-day flowering in rice. Nature 422 (6933):719-722. doi:DOI 10.1038/nature01549

He Y (2012) Chromatin regulation of flowering. Trends in Plant Science 17 (9):556-562. doi:10.1016/j.tplants.2012.05.001

He Y, Doyle MR, Amasino RM (2004) PAF1-complex-mediated histone methylation of FLOWERING LOCUS C chromatin is required for the vernalization-responsive, winter-annual habit in Arabidopsis. Genes and Development 18 (22):2774-2784. doi:10.1101/gad.1244504

He YH, Michaels SD, Amasino RM (2003) Regulation of flowering time by histone acetylation in Arabidopsis. Science 302 (5651):1751-1754. doi:DOI 10.1126/science.1091109

Hecht V, Foucher F, Ferrandiz C, Macknight R, Navarro C, Morin J, Vardy ME, Ellis N, Beltran JP, Rameau C, Weller JL (2005) Conservation of Arabidopsis flowering genes in model legumes. Plant Physiology 137 (4):1420-1434. doi:10.1104/pp.104.057018

Hecht V, Knowles CL, Schoor JKV, Liew LC, Jones SE, Lambert MJM, Weller JL (2007) Pea LATE BLOOMER1 is a GIGANTEA ortholog with roles in photoperiodic flowering, deetiolation, and transcriptional regulation of circadian clock gene homologs. Plant Physiology 144 (2):648-661. doi:10.1104/pp.107.096818

Hecht V, Laurie RE, Vander Schoor JK, Ridge S, Knowles CL, Liew LC, Sussmilch FC, Murfet IC, Macknight RC, Weller JL (2011) The pea GIGAS gene is a FLOWERING LOCUS T homolog necessary for graft-transmissible specification of flowering but not for responsiveness to photoperiod. The Plant Cell 23 (1):147-161. doi:10.1105/tpc.110.081042

Hectors K, Prinsen E, De Coen W, Jansen MA, Guisez Y (2007) Arabidopsis thaliana plants acclimated to low dose rates of ultraviolet B radiation show specific changes in morphology and gene expression in the absence of stress symptoms. The New Phytologist 175 (2):255-270. doi:10.1111/j.1469-8137.2007.02092.x

Heijde M, Ulm R (2012) UV-B photoreceptor-mediated signalling in plants. Trends in Plant Science 17 (4):230-237. doi:10.1016/j.tplants.2012.01.007

Hejatko J, Ryu H, Kim GT, Dobesova R, Choi S, Choi SM, Soucek P, Horak J, Pekarova B, Palme K, Brzobohaty B, Hwang I (2009) The Histidine Kinases CYTOKININ-INDEPENDENT1 and ARABIDOPSIS HISTIDINE KINASE2 and 3 Regulate Vascular Tissue Development in Arabidopsis Shoots. The Plant Cell 21 (7):2008-2021. doi:10.1105/tpc.109.066696

Hemsley PA, Hurst CH, Kaliyadasa E, Lamb R, Knight MR, De Cothi EA, Steele JF, Knight H (2014) The Arabidopsis Mediator Complex Subunits MED16, MED14, and MED2 Regulate Mediator and RNA Polymerase II Recruitment to CBF-Responsive Cold-Regulated Genes. The Plant Cell 26 (1):465-484. doi:10.1105/tpc.113.117796

Henikoff S, Henikoff JG (1992) Amino acid substitution matrices from protein blocks. Proceedings of the National Academy of Sciences 89 (22):10915-10919

Henry IM, Nagalakshmi U, Lieberman MC, Ngo KJ, Krasileva KV, Vasquez-Gross H, Akhunova A, Akhunov E, Dubcovsky J, Tai TH, Comai L (2014) Efficient Genome-Wide Detection and Cataloging of EMS-Induced Mutations Using Exome Capture and Next-Generation Sequencing. The Plant cell 26 (4):1382-1397. doi:10.1105/tpc.113.121590

Heo JB, Sung S (2011) Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. Science (New York, NY) 331 (6013):76-79. doi:10.1126/science.1197349

Hernandez J, Olmos E, Corpas F, Sevilla F, Del Rio L (1995) Salt-induced oxidative stress in chloroplasts of pea plants. Plant Science 105 (2):151-167

Heuer S, Hansen S, Bantin J, Brettschneider R, Kranz E, Lörz H, Dresselhaus T (2001) The maize MADS box gene ZmMADS3 affects node number and spikelet development and is co-expressed with ZmMADS1 during flower development, in egg cells, and early embryogenesis. Plant Physiology 127 (1):33-45. doi:10.1104/pp.127.1.33

Higgins DG, Thompson JD, Gibson TJ (1996) [22] Using CLUSTAL for multiple sequence alignments. In: Methods in Enzymology, vol 266. Academic Press, pp 383-402. doi:https://doi.org/10.1016/S0076-6879(96)66024-8

Higgins JA, Bailey PC, Laurie DA (2010) Comparative genomics of flowering time pathways using Brachypodium distachyon as a model for the temperate grasses. PLOS One 5 (4):e10065

Highkin HR (1956) Vernalization in Peas. Plant Physiology 31 (5):399-403. doi:10.1104/pp.31.5.399

Hill CB, Li C (2016) Genetic Architecture of Flowering Phenology in Cereals and Opportunities for Crop Improvement. Frontiers in Plant Science 7:1906. doi:10.3389/fpls.2016.01906

Hofer J, Turner L, Hellens R, Ambrose M, Matthews P, Michael A, Ellis N (1997) UNIFOLIATA regulates leaf and flower morphogenesis in pea. Current Biology 7 (8):581-587. doi:https://doi.org/10.1016/S0960-9822(06)00257-0

Hofer J, Turner L, Moreau C, Ambrose M, Isaac P, Butcher S, Weller J, Dupin A, Dalmais M, Le Signor C, Bendahmane A, Ellis N (2009) Tendril-less Regulates Tendril Formation in Pea Leaves. The Plant Cell 21 (2):420-428. doi:10.1105/tpc.108.064071

Holstege FCP, Jennings EG, Wyrick JJ, Lee TI, Hengartner CJ, Green MR, Golub TR, Lander ES, Young RA (1998) Dissecting the regulatory circuitry of a eukaryotic genome. Cell 95 (5):717-728. doi:Doi 10.1016/S0092-8674(00)81641-4

Hong Y, Jackson S (2015) Floral induction and flower formation--the role and potential applications of miRNAs. Plant Biotechnology Journal 13 (3):282-292. doi:10.1111/pbi.12340

Hornyik C, Terzi LC, Simpson GG (2010) The spen family protein FPA controls alternative cleavage and polyadenylation of RNA. Developmental Cell 18 (2):203-213. doi:10.1016/j.devcel.2009.12.009

Hsu PY, Harmer SL (2014) Wheels within wheels: the plant circadian system. Trends in Plant Science 19 (4):240-249. doi:10.1016/j.tplants.2013.11.007

Huang LS, Sternberg PW (1995) Genetic dissection of developmental pathways. Methods in Cell Biology 48:97-122

Huang LS, Sternberg PW (2006) Genetic dissection of developmental pathways. WormBook:1-19. doi:10.1895/wormbook.1.88.2

Huang S, Raman AS, Ream JE, Fujiwara H, Cerny RE, Brown SM (1998) Overexpression of 20-oxidase confers a gibberellin-overproduction phenotype in Arabidopsis. Plant Physiology 118 (3):773-781

Huang W, Perez-Garcia P, Pokhilko A, Millar AJ, Antoshechkin I, Riechmann JL, Mas P (2012) Mapping the Core of the Arabidopsis Circadian Clock Defines the Network Structure of the Oscillator. Science 336 (6077):75-79. doi:DOI 10.1126/science.1219075

Huijser P, Schmid M (2011) The control of developmental phase transitions in plants. Development 138 (19):4117-4129. doi:DOI 10.1242/dev.063511

Hung HY, Shannon LM, Tian F, Bradbury PJ, Chen C, Flint-Garcia SA, McMullen MD, Ware D, Buckler ES, Doebley JF, Holland JB (2012) ZmCCT and the genetic basis of day-length adaptation underlying the postdomestication spread of maize. Proceedings of the National Academy of Sciences of the United States of America 109 (28):E1913-E1921. doi:10.1073/pnas.1203189109

Ietswaart R, Wu Z, Dean C Flowering time control: another window to the connection between antisense RNA and chromatin. Trends in Genetics 28 (9):445-453. doi:10.1016/j.tig.2012.06.002

Imura Y, Kobayashi Y, Yamamoto S, Furutani M, Tasaka M, Abe M, Araki T (2012) CRYPTIC PRECOCIOUS/MED12 is a Novel Flowering Regulator with Multiple Target Steps in Arabidopsis. Plant and Cell Physiology 53 (2):287-303. doi:10.1093/pcp/pcs002

Iñigo S, Alvarez MJ, Strasser B, Califano A, Cerdán PD (2012) PFT1, the MED25 subunit of the plant Mediator complex, promotes flowering through CONSTANS dependent and independent mechanisms in Arabidopsis. The Plant Journal 69 (4):601-612

Inzé D (2007) The cell cycle control and plant development. Annual plant reviews: v. 32. Oxford, UK ; Ames, Iowa : Blackwell Pub., 2007.,

Ito J, Sono T, Tasaka M, Furutani M (2011) MACCHI-BOU 2 is Required for Early Embryo Patterning and Cotyledon Organogenesis in Arabidopsis. Plant and Cell Physiology 52 (3):539-552. doi:10.1093/pcp/pcr013

Iwata H, Gotoh O (2011) Comparative analysis of information contents relevant to recognition of introns in many species. BMC Genomics 12. doi:Artn 4510.1186/1471-2164-12-45

Izaguirre MM, Mazza CA, Biondini M, Baldwin IT, Ballare CL (2006) Remote sensing of future competitors: Impacts on plant defenses. Proceedings of the National Academy of Sciences of the United States of America 103 (18):7170-7174. doi:10.1073/pnas.0509805103

Jacinto T, McGurl B, Franceschi V, DelanoFreier J, Ryan CA (1997) Tomato prosystemin promoter confers wound-inducible, vascular bundle-specific expression of the beta-glucuronidase gene in transgenic tomato plants. Planta 203 (4):406-412. doi:DOI 10.1007/s004250050207

Jackson JP, Lindroth AM, Cao XF, Jacobsen SE (2002) Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature 416 (6880):556-560. doi:DOI 10.1038/nature731

Jacob Y, Feng S, LeBlanc CA, Bernatavichute YV, Stroud H, Cokus S, Johnson LM, Pellegrini M, Jacobsen SE, Michaels SD (2009) ATXR5 and ATXR6 are H3K27 monomethyltransferases required for chromatin structure and gene silencing. Nature Structural &Amp; Molecular Biology 16:763. doi:10.1038/nsmb.1611

Jang S, Marchal V, Panigrahi KCS, Wenkel S, Soppe W, Deng XW, Valverde F, Coupland G (2008) Arabidopsis COP1 shapes the temporal pattern of CO accumulation conferring a photoperiodic flowering response. EMBO Journal 27 (8):1277-1288. doi:DOI 10.1038/emboj.2008.68

Jang S, Torti S, Coupland G (2009) Genetic and spatial interactions between FT, TSF and SVP during the early stages of floral induction in Arabidopsis. The Plant Journal : for Cell and Molecular Biology 60 (4):614-625. doi:10.1111/j.1365-313X.2009.03986.x

Janody F, Martirosyan Z, Benlali A, Treisman JE (2003) Two subunits of the Drosophila mediator complex act together to control cell affinity. Development 130 (16):3691-3701. doi:DOI 10.1242/dev.00607

Jansen MAK (2002) Ultraviolet-B radiation effects on plants: induction of morphogenic responses. Physiologia Plantarum 116 (3):423-429. doi:DOI 10.1034/j.1399-3054.2002.1160319.x

Jarvik J, Botstein D (1973) A genetic method for determining the order of events in a biological pathway. Proceedings of the National Academy of Sciences of the United States of America 70 (7):2046-2050

Jayasinghege CPA, Ozga JA, Waduthanthri KD, Reinecke DM (2017) Regulation of ethylene-related gene expression by indole-3-acetic acid and 4-chloroindole-3-acetic acid in relation to pea fruit and seed development. Journal of Experimental Botany 68 (15):4137-4151. doi:10.1093/jxb/erx217

Jeronimo C, Robert F (2017) The Mediator Complex: At the Nexus of RNA Polymerase II Transcription. Trends in Cell Biology 27 (10):765-783. doi:https://doi.org/10.1016/j.tcb.2017.07.001

Jiang D, Kong NC, Gu X, Li Z, He Y (2011) Arabidopsis COMPASS-like complexes mediate histone H3 lysine-4 trimethylation to control floral transition and plant development. PLOS Genetics 7 (3). doi:10.1371/journal.pgen.1001330

Jiang D, Wang Y, Wang Y, He Y (2008) Repression of FLOWERING LOCUS C and FLOWERING LOCUS T by the Arabidopsis Polycomb Repressive Complex 2 Components. PLOS One 3 (10). doi:10.1371/journal.pone.0003404

Jin B, Wang L, Wang J, Jiang KZ, Wang Y, Jiang XX, Ni CY, Wang YL, Teng NJ (2011) The effect of experimental warming on leaf functional traits, leaf structure and leaf biochemistry in Arabidopsis thaliana. BMC Plant Biology 11. doi:Artn 3510.1186/1471-2229-11-35

Jin J, Tian F, Yang D-C, Meng Y-Q, Kong L, Luo J, Gao G (2017) PlantTFDB 4.0: toward a central hub for transcription factors and regulatory interactions in plants. Nucleic Acids Research 45 (D1):D1040-D1045. doi:10.1093/nar/gkw982

Jin J, Zhang H, Kong L, Gao G, Luo J (2014) PlantTFDB 3.0: a portal for the functional and evolutionary study of plant transcription factors. Nucleic Acids Research 42 (D1):D1182-D1187. doi:10.1093/nar/gkt1016

Jing R, Johnson R, Seres A, Kiss G, Ambrose MJ, Knox MR, Ellis THN, Flavell AJ (2007) Gene-based sequence diversity analysis of field pea (Pisum). Genetics 177 (4):2263-2275. doi:10.1534/genetics107.081323

Johansson M, Staiger D (2015) Time to flower: interplay between photoperiod and the circadian clock. Journal of Experimental Botany 66 (3):719-730. doi:10.1093/jxb/eru441

Johnson X, Brcich T, Dun EA, Goussot M, Haurogné K, Beveridge CA, Rameau C (2006) Branching genes are conserved across species. Genes controlling a novel signal in pea are coregulated by other long-distance signals. Plant Physiology 142 (3):1014-1026

Jones S, Thornton JM (1996) Principles of protein-protein interactions. Proceedings of the National Academy of Sciences of the United States of America 93 (1):13-20

Josse E-M, Halliday KJ (2008) Skotomorphogenesis: The Dark Side of Light Signalling. Current Biology 18 (24):R1144-R1146. doi:https://doi.org/10.1016/j.cub.2008.10.034

Kaló P, Seres A, Taylor SA, Jakab J, Kevei Z, Kereszt A, Endre G, Ellis THN, Kiss GB (2004) Comparative mapping between Medicago sativa and Pisum sativum. Molecular Genetics and Genomics 272 (3):235-246. doi:10.1007/s00438-004-1055-z

Kami C, Lorrain S, Hornitschek P, Fankhauser C (2010) Light-Regulated Plant Growth and Development. Plant Development 91:29-66. doi:10.1016/S0070-2153(10)91002-8

Kaneko-Suzuki M, Kurihara-Ishikawa R, Okushita-Terakawa C, Kojima C, Nagano-Fujiwara M, Ohki I, Tsuji H, Shimamoto K, Taoka K-I (2018) TFL1-Like Proteins in Rice Antagonize Rice FT-Like Protein in Inflorescence Development by Competition for Complex Formation with 14-3-3 and FD. Plant and Cell Physiology 59 (3):458-468. doi:10.1093/pcp/pcy021

Kaur S, Pembleton LW, Cogan NOI, Savin KW, Leonforte T, Paull J, Materne M, Forster JW (2012) Transcriptome sequencing of field pea and faba bean for discovery and validation of SSR genetic markers. BMC Genomics 13:104-104. doi:10.1186/1471-2164-13-104

Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28 (12):1647-1649. doi:10.1093/bioinformatics/bts199

Kelleher RJ, Flanagan PM, Kornberg RD (1990) A Novel Mediator between Activator Proteins and the Rna Polymerase-Ii Transcription Apparatus. Cell 61 (7):1209-1215. doi:Doi 10.1016/0092-8674(90)90685-8

Khan M, Rozhon W, Poppenberger B (2014) The role of hormones in the aging of plants-a mini-review. Gerontology 60 (1):49-55

Kidd BN, Cahill DM, Manners JM, Schenk PM, Kazan K (2011) Diverse roles of the Mediator complex in plants. Seminars in Cell & Developmental Biology 22 (7):741-748. doi:https://doi.org/10.1016/j.semcdb.2011.07.012

Kidd BN, Edgar CI, Kumar KK, Aitken EA, Schenk PM, Manners JM, Kazan K (2009) The Mediator Complex Subunit PFT1 Is a Key Regulator of Jasmonate-Dependent Defense in Arabidopsis. The Plant Cell 21 (8):2237-2252. doi:10.1105/tpc.109.066910

Kilian J, Whitehead D, Horak J, Wanke D, Weinl S, Batistic O, D'Angelo C, Bornberg-Bauer E, Kudla J, Harter K (2007) The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. The Plant Journal : for Cell and Molecular Biology 50 (2):347-363. doi:10.1111/j.1365-313X.2007.03052.x

Kim DH, Doyle MR, Sung S, Amasino RM (2009) Vernalization: Winter and the timing of flowering in plants. Annual Review of Cell and Developmental Biology, vol 25. doi:10.1146/annurev.cellbio.042308.113411

Kim JJ, Lee JH, Kim W, Jung HS, Huijser P, Ahn JH (2012) The microRNA156-SQUAMOSA PROMOTER BINDING PROTEIN-LIKE3 Module Regulates Ambient Temperature-Responsive Flowering via FLOWERING LOCUS T in Arabidopsis. Plant Physiology 159 (1):461-478. doi:DOI 10.1104/pp.111.192369

Kim S, Choi K, Park C, Hwang HJ, Lee I (2006) SUPPRESSOR of FRIGIDA4, encoding a C2H2-type zinc finger protein, represses flowering by transcriptional activation of Arabidopsis FLOWERING LOCUS C. The Plant Cell 18 (11):2985-2998. doi:10.1105/tpc.106.045179

Kim SL, Lee S, Kim HJ, Nam HG, An G (2007) OsMADS51 is a short-day flowering promoter that functions upstream of Ehd1, OsMADS14, and Hd3a. Plant Physiology 145 (4):1484-1494. doi:10.1104/pp.107.103291

Kim TW, Kwon Y-J, Kim JM, Song Y-H, Kim SN, Kim Y-J (2004) MED16 and MED23 of Mediator are coactivators of lipopolysaccharide- and heat-shock-induced transcriptional activators. Proceedings of the National Academy of Sciences of the United States of America 101 (33):12153-12158. doi:10.1073/pnas.0401985101

Klem K, Ac A, Holub P, Kovac D, Spunda V, Robson TM, Urban O (2012) Interactive effects of PAR and UV radiation on the physiology, morphology and leaf optical properties of two barley varieties. Environmental and Experimental Botany75:52-64. doi:10.1016/j.envexpbot.2011.08.008

Kliebenstein DJ, Lim JE, Landry LG, Last RL (2002) Arabidopsis UVR8 regulates ultraviolet-B signal transduction and tolerance and contains sequence similarity to human Regulator of Chromatin Condensation 1. Plant Physiology 130 (1):234-243. doi:10.1104/pp.005041

Knuesel MT, Meyer KD, Bernecky C, Taatjes DJ (2009a) The human CDK8 subcomplex is a molecular switch that controls Mediator coactivator function. Genes & Development23 (4):439-451. doi:10.1101/gad.1767009

Knuesel MT, Meyer KD, Donner AJ, Espinosa JM, Taatjes DJ (2009b) The Human CDK8 Subcomplex Is a Histone Kinase That Requires Med12 for Activity and Can Function Independently of Mediator. Molecular and Cellular Biology 29 (3):650-661. doi:10.1128/MCB.00993-08

Ko JH, Mitina I, Tamada Y, Hyun Y, Choi Y, Amasino RM, Noh B, Noh YS (2010) Growth habit determination by the balance of histone methylation activities in Arabidopsis. EMBO Journal 29 (18):3208-3215. doi:10.1038/emboj.2010.198

Kojima S, Takahashi Y, Kobayashi Y, Monna L, Sasaki T, Araki T, Yano M (2002) Hd3a, a rice ortholog of the Arabidopsis FT gene, promotes transition to flowering downstream of Hd1 under short-day conditions. Plant and Cell Physiology 43 (10):1096-1105. doi:DOI 10.1093/pcp/pcf156

Kong F, Nan H, Cao D, Li Y, Wu F, Wang J, Lu S, Yuan X, Cober ER, Abe J, Liu B (2014) A new dominant gene E9 conditions early flowering and maturity in soybean. Crop Science 54 (6):2529-2535. doi:10.2135/cropsci2014.03.0228

Kozak M (1980) Evaluation of the Scanning Model for Initiation of Protein-Synthesis in Eukaryotes. Cell 22 (1):7-8. doi:Doi 10.1016/0092-8674(80)90148-8

Kozak M (1984a) Compilation and Analysis of Sequences Upstream from the Translational Start Site in Eukaryotic Messenger-Rnas. Nucleic Acids Research 12 (2):857-872. doi:DOI 10.1093/nar/12.2.857

Kozak M (1984b) Point Mutations Close to the Aug Initiator Codon Affect the Efficiency of Translation of Rat Preproinsulin Invivo. Nature 308 (5956):241-246. doi:DOI 10.1038/308241a0

Kozak M (1984c) Selection of Initiation Sites by Eukaryotic Ribosomes - Effect of Inserting Aug Triplets Upstream from the Coding Sequence for Preproinsulin. Nucleic Acids Research 12 (9):3873-3893. doi:DOI 10.1093/nar/12.9.3873

Kozak M (1987) An Analysis of 5'-Noncoding Sequences from 699 Vertebrate Messenger-Rnas. Nucleic Acids Research 15 (20):8125-8148. doi:DOI 10.1093/nar/15.20.8125

Kozak M (1989) The Scanning Model for Translation - an Update. Journal of Cell Biology108 (2):229-241. doi:DOI 10.1083/jcb.108.2.229

Kozak M (1992) A Consideration of Alternative Models for the Initiation of Translation in Eukaryotes. Critical Reviews in Biochemistry and Molecular Biology27 (4-5):385-402. doi:Doi 10.3109/10409239209082567

Kozak M (1995) Adherence to the First-Aug Rule When a 2nd-Aug Codon Follows Closely Upon the First (Vol 92, Pg 2662, 1995). Proceedings of the National Academy of Sciences of the United States of America 92 (15):7134-7134. doi:DOI 10.1073/pnas.92.15.7134

Kozak M (2002) Emerging links between initiation of translation and human diseases. Mammalian Genome 13 (8):401-410. doi:10.1007/s00335-002-4002-5

Krasensky J, Jonak C (2012) Drought, salt, and temperature stress-induced metabolic rearrangements and regulatory networks. Journal of Experimental Botany 63 (4):1593-1608. doi:10.1093/jxb/err460

Kronzucker HJ, Coskun D, Schulze LM, Wong JR, Britto DT (2013) Sodium as nutrient and toxicant. Plant Soil 369 (1-2):1-23. doi:10.1007/s11104-013-1801-2

Kumar SV, Lucyshyn D, Jaeger KE, Alos E, Alvey E, Harberd NP, Wigge PA (2012) Transcription factor PIF4 controls the thermosensory activation of flowering. Nature 484 (7393):242-245. doi:10.1038/nature10928

Kumar SV, Wigge PA (2010) H2A.Z-containing nucleosomes mediate the thermosensory response in Arabidopsis. Cell 140 (1):136-147. doi:10.1016/j.cell.2009.11.006

Kumar V (2013) Systems Biology Approaches towards the Prediction of Prospective Novel Plant System-Derived Products or Services. Journal of Computer Science and Systems Biology 2 (4). doi:10.4172/2329-6577.1000119

Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG (2007) Clustal W and Clustal X version 2.0. Bioinformatics 23 (21):2947-2948. doi:10.1093/bioinformatics/btm404

Larschan E, Winston F (2005) The Saccharomyces cerevisiae Srb8-Srb11 complex functions with the SAGA complex during Gal4-activated transcription. Molecular and Cellular Biology 25 (1):114-123. doi:10.1128/Mcb.25.1.114-123.2005

Laubinger S, Marchal V, Gentilhomme J, Wenkel S, Adrian J, Jang S, Kulajta C, Braun H, Coupland G, Hoecker U (2006) Arabidopsis SPA proteins regulate photoperiodic flowering and interact with the floral inducer CONSTANS to regulate its stability (vol 133, pg 3213, 2006). Development 133 (22):4608-4608. doi:DOI 10.1242/dev.02691

Laurie RE, Diwadkar P, Jaudal M, Zhang L, Hecht V, Wen J, Tadege M, Mysore KS, Putterill J, Weller JL, Macknight RC (2011) The Medicago FLOWERING LOCUS T homolog, MtFTa1, is a key regulator of flowering time. Plant Physiology 156 (4):2207-2224. doi:10.1104/pp.111.180182

Lee I, Ambaru B, Thakkar P, Marcotte EM, Rhee SY (2010) Rational association of genes with traits using a genome-scale gene network for Arabidopsis thaliana. Nature Biotechnology 28 (2):149-156. doi:10.1038/nbt.1603

Lee JH, Yoo SJ, Park SH, Hwang I, Lee JS, Ahn JH (2007) Role of SVP in the control of flowering time by ambient temperature in Arabidopsis. Genes & Development21 (4):397-402. doi:DOI 10.1101/gad.1518407

Lee T, Lee I (2017) AraNet: A Network Biology Server for Arabidopsis thaliana and Other Non-Model Plant Species. In: Kaufmann K, Mueller-Roeber B (eds) Plant Gene Regulatory Networks: Methods and Protocols. Springer New York, New York, NY, pp 225-238. doi:10.1007/978-1-4939-7125-1_15

Lee T, Yang S, Kim E, Ko Y, Hwang S, Shin J, Shim JE, Shim H, Kim H, Kim C, Lee I (2015) AraNet v2: an improved database of co-functional gene networks for the study of Arabidopsis thaliana and 27 other nonmodel plant species. Nucleic Acids Research 43 (Database issue):D996-D1002. doi:10.1093/nar/gku1053

Lens F, Smets E, Melzer S (2012) Stem anatomy supports Arabidopsis thaliana as a model for insular woodiness. New Phytologist 193 (1):12-17

León J, Rojo E, Sánchez‐Serrano JJ (2001) Wound signalling in plants. Journal of Experimental Botany 52 (354):1-9. doi:10.1093/jexbot/52.354.1

Leonforte A, Forster JW, Redden RJ, Nicolas ME, Salisbury PA (2013) Sources of high tolerance to salinity in pea (Pisum sativum L.). Euphytica 189 (2):203-216. doi:10.1007/s10681-012-0771-4

Leshchiner I, Alexa K, Kelsey P, Adzhubei I, Austin-Tse CA, Cooney JD, Anderson H, King MJ, Stottmann RW, Garnaas MK, Ha SS, Drummond IA, Paw BH, North TE, Beier DR, Goessling W, Sunyaev SR (2012) Mutation mapping and identification by whole-genome sequencing. Genome Research 22 (8):1541-1548. doi:10.1101/gr.135541.111

Lesser C, Guthrie C (1993) Mutations in U6 snRNA that alter splice site specificity: implications for the active site. Science 262 (5142):1982-1988. doi:10.1126/science.8266093

Lester DR, MacKenzie-Hose AK, Davies PJ, Ross JJ, Reid JB (1999) The influence of the null le-2 mutation on gibberellin levels in developing pea seeds. Journal of Plant Growth Regulation 27 (2):83-89. doi:Doi 10.1023/A:1006158701016

Lewandowska MA (2013) The missing puzzle piece: splicing mutations. International Journal of Clinical and Experimental Pathology 6 (12):2675-2682

Lewandowski TJ, Dunfield KE, Antunes PM (2013) Isolate Identity Determines Plant Tolerance to Pathogen Attack in Assembled Mycorrhizal Communities. PLOS One 8 (4):e61329. doi:10.1371/journal.pone.0061329

Li B, Wachtel C, Miriami E, Yahalom G, Friedlander G, Sharon G, Sperling R, Sperling J (2002) Stop codons affect 5′ splice site selection by surveillance of splicing. Proceedings of the National Academy of Sciences of the United States of America 99 (8):5277-5282. doi:10.1073/pnas.082095299

Li J, Han Y, Zhao Q, Li C, Xie Q, Chong K, Xu Y (2013a) The E3 Ligase AtRDUF1 Positively Regulates Salt Stress Responses in Arabidopsis thaliana. PLOS One 8 (8):e71078. doi:10.1371/journal.pone.0071078

Li L, Li D, Liu S, Ma X, Dietrich CR, Hu H-C, Zhang G, Liu Z, Zheng J, Wang G, Schnable PS (2013b) The Maize glossy13 Gene, Cloned via BSR-Seq and Seq-Walking Encodes a Putative ABC Transporter Required for the Normal Accumulation of Epicuticular Waxes. PLOS One 8 (12):e82333. doi:10.1371/journal.pone.0082333

Li Z, Wakao S, Fischer BB, Niyogi KK (2009) Sensing and responding to excess light. Annual Review of Plant Biology 60:239-260. doi:10.1146/annurev.arplant.58.032806.103844

Liew LC, Hecht V, Laurie RE, Knowles CL, Vander Schoor JK, Macknight RC, Weller JL (2009) DIE NEUTRALIS and LATE BLOOMER 1 contribute to regulation of the pea circadian clock. The Plant Cell 21 (10):3198-3211. doi:10.1105/tpc.109.067223

Liew LC, Hecht V, Sussmilch FC, Weller JL (2014) The Pea Photoperiod Response Gene STERILE NODES Is an Ortholog of LUX ARRHYTHMO. Plant Physiology 165 (2):648-657. doi:10.1104/pp.114.237008

Ligerot Y, de Saint Germain A, Waldie T, Troadec C, Citerne S, Kadakia N, Pillot J-P, Prigge M, Aubert G, Bendahmane A, Leyser O, Estelle M, Debellé F, Rameau C (2017) The pea branching RMS2 gene encodes the PsAFB4/5 auxin receptor and is involved in an auxin-strigolactone regulation loop. PLOS Genetics 13 (12):e1007089. doi:10.1371/journal.pgen.1007089

Liljegren SJ, Gustafson-Brown C, Pinyopich A, Ditta GS, Yanofsky MF (1999) Interactions among APETALA1, LEAFY, and TERMINAL FLOWER1 specify meristem fate. The Plant Cell 11 (6):1007-1018. doi:10.1105/tpc.11.6.1007

Lim LP, Burge CB (2001) A computational analysis of sequence features involved in recognition of short introns. Proceedings of the National Academy of Sciences 98 (20):11193-11198. doi:10.1073/pnas.201407298

Lincoln SE, Daly MJ, Lander ES (1993) Constructing genetic linkage maps with MAPMAKER/EXP Version 3.0: a tutorial and reference manual. A whitehead institute for biomedical research technical report 3

Lindroth AM, Shultis D, Jasencakova Z, Fuchs J, Johnson L, Schubert D, Patnaik D, Pradhan S, Goodrich J, Schubert I, Jenuwein T, Khorasanizadeh S, Jacobsen SE (2004) Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with CHROMOMETHYLASE3. EMBO Journal 23 (21):4286-4296. doi:10.1038/sj.emboj.7600430

Liu B, Asseng S, Wang A, Wang S, Tang L, Cao W, Zhu Y, Liu L (2017) Modelling the effects of post-heading heat stress on biomass growth of winter wheat. Agricultural and Forest Meteorology 247:476-490. doi:https://doi.org/10.1016/j.agrformet.2017.08.018

Liu CY, Lu FL, Cui X, Cao XF (2010) Histone Methylation in Higher Plants. Annual Review of Plant Biology, Vol 61 61:395-420. doi:10.1146/annurev.arplant.043008.091939

Liu F, Quesada V, Crevillen P, Baurle I, Swiezewski S, Dean C (2007) The Arabidopsis RNA-binding protein FCA requires a lysine-specific demethylase 1 homolog to downregulate FLC. Molecular Cell 28 (3):398-407. doi:10.1016/j.molcel.2007.10.018

Liu LJ, Zhang YC, Li QH, Sang Y, Mao J, Lian HL, Wang L, Yang HQ (2008) COP1-mediated ubiquitination of CONSTANS is implicated in cryptochrome regulation of flowering in Arabidopsis. The Plant Cell 20 (2):292-306. doi:10.1105/tpc.107.057281

Liu S, Yeh C-T, Tang HM, Nettleton D, Schnable PS (2012) Gene mapping via bulked segregant RNA-Seq (BSR-Seq). PLOS One 7 (5):e36406

Liu Z, Mara C (2010) Regulatory mechanisms for floral homeotic gene expression. Semin Cell Dev Biol 21 (1):80-86. doi:10.1016/j.semcdb.2009.11.012

Llugany M, Martin SR, Barceló J, Poschenrieder C (2013) Endogenous jasmonic and salicylic acids levels in the Cd-hyperaccumulator Noccaea (Thlaspi) praecox exposed to fungal infection and/or mechanical stress. Plant Cell Reports 32 (8):1243-1249. doi:10.1007/s00299-013-1427-0

Lockhart JA, Gottschall V (1961) Fruit-induced & apical senescence in Pisum sativum L. Plant Physiology 36 (4):389

Loncle N, Boube M, Joulia L, Boschiero C, Werner M, Cribbs DL, Bourbon H-M (2007) Distinct roles for Mediator Cdk8 module subunits in Drosophila development. The EMBO Journal 26 (4):1045-1054. doi:10.1038/sj.emboj.7601566

Long Jennifer C, Caceres Javier F (2009) The SR protein family of splicing factors: master regulators of gene expression. Biochemical Journal 417 (1):15-27. doi:10.1042/bj20081501

Loridon K, McPhee K, Morin J, Dubreuil P, Pilet-Nayel ML, Aubert G, Rameau C, Baranger A, Coyne C, Lejeune-Hènaut I, Burstin J (2005) Microsatellite marker polymorphism and mapping in pea (Pisum sativum L.). Theoretical and Applied Genetics 111 (6):1022-1031. doi:10.1007/s00122-005-0014-3

Love J, Bjorklund S, Vahala J, Hertzberg M, Kangasjarvi J, Sundberg B (2009) Ethylene is an endogenous stimulator of cell division in the cambial meristem of Populus. Proceedings of the National Academy of Sciences of the United States of America 106 (14):5984-5989. doi:10.1073/pnas.0811660106

Lovett PS, Rogers EJ (1996) Ribosome regulation by the nascent peptide. Microbiological Reviews 60 (2):366-385

Lozano F, Maertzdorf B, Pannell R, Milstein C (1994) Low cytoplasmic mRNA levels of immunoglobulin kappa light chain genes containing nonsense codons correlate with inefficient splicing. The EMBO Journal 13 (19):4617-4622

Lu SJ, Zhao XH, Hu YL, Liu SL, Nan HY, Li XM, Fang C, Cao D, Shi XY, Kong LP, Su T, Zhang FG, Li SC, Wang Z, Yuan XH, Cober ER, Weller JL, Liu BH, Hou XL, Tian ZX, Kong FJ (2017) Natural variation at the soybean J locus improves adaptation to the tropics and enhances yield. Nature Genetics 49 (5):773-+. doi:10.1038/ng.3819

Luo M, Tai R, Yu CW, Yang S, Chen CY, Lin WD, Schmidt W, Wu K (2015) Regulation of flowering time by the histone deacetylase HDA5 in Arabidopsis. The Plant Journal : for Cell and Molecular Biology 82 (6):925-936. doi:10.1111/tpj.12868

Lykke-Andersen S, Jensen TH (2015) Nonsense-mediated mRNA decay: an intricate machinery that shapes transcriptomes. Nature Reviews Molecular Cell Biology 16:665. doi:10.1038/nrm4063

Macas J, Neumann P, Navrátilová A (2007) Repetitive DNA in the pea (Pisum sativum L.) genome: comprehensive characterization using 454 sequencing and comparison to soybean and Medicago truncatula. BMC Genomics 8 (1):427. doi:10.1186/1471-2164-8-427

Madoui MA KL, JM Aury, CJ Coyne, T Warkentin, J Kreplak (2015) The international pea genome sequencing project : sequencing and assembly progresses PLANT and ANIMAL Genome XXIII (SanDiego, CA)

Mahony J (1991) Field pea. *In* New Crops: Agronomy and Potential of Alternative Crop Species Eds RS Jessop and RL Wright:53-66. Inkata Press, Melbourne

Malik N, Agarwal P, Tyagi A (2017) Emerging functions of multi-protein complex Mediator with special emphasis on plants. Critical Reviews in Biochemistry and Molecular Biology52 (5):475-502. doi:10.1080/10409238.2017.1325830

Manly KF, Cudmore RH, Meer JM (2001) Map manager QTX, cross-platform software for genetic mapping. Mammalian Genome 12 (12):930-932. doi:DOI 10.1007/s00335-001-1016-3

Maquat LE (1995) When cells stop making sense: effects of nonsense codons on RNA metabolism in vertebrate cells. RNA 1 (5):453-465

Maquat LE (2005) Nonsense-mediated mRNA decay in mammals. Journal of Cell Science 118 (9):1773-1776. doi:10.1242/jcs.01701

Marschner H (2011) Marschner's mineral nutrition of higher plants. Academic press,

Martin C, Zhang Y (2005) The diverse functions of histone lysine methylation. Nature Reviews Molecular Cell Biology 6 (11):838-849. doi:DOI 10.1038/nrm1761

Masaya I, Mikiko K, Hitoshi S, Shoko K, Yoshihisa U, Chiyoko M, Yasunori M (2010) Genetic networks regulated by ASYMMETRIC LEAVES1 (AS1) and AS2 in leaf development in Arabidopsis thaliana: KNOX genes control five morphological events. The Plant Journal 61 (1):70-82. doi:doi:10.1111/j.1365-313X.2009.04033.x

Mascheretti I, Turner K, Brivio RS, Hand A, Colasanti J, Rossi V (2015) Focus on chromatin/epigenetics: florigen-encoding genes of day-neutral and photoperiod-sensitive maize are regulated by different chromatin modifications at the floral transition. Plant Physiology 168 (4):1351

Mathieu J, Yant LJ, Murdter F, Kuttner F, Schmid M (2009) Repression of Flowering by the miR172 Target SMZ. PLOS Biology 7 (7). doi:ARTN e1000148

DOI 10.1371/journal.pbio.1000148

Mathur S, Vyas S, Kapoor S, Tyagi AK (2011) The Mediator Complex in Plants: Structure, Phylogeny, and Expression Profiling of Representative Genes in a Dicot (Arabidopsis) and a Monocot (Rice) during Reproduction and Abiotic Stress. Plant Physiology 157 (4):1609-1627. doi:10.1104/pp.111.188300

McAdam SAM, Brodribb TJ, Ross JJ (2016) Shoot-derived abscisic acid promotes root growth. Plant Cell and Environment 39 (3):652-659. doi:10.1111/pce.12669

McCall D, Atherton JG (1995) Interactions between diurnal temperature fluctuations and salinity on expansion growth and water status of young tomato plants. Annals of Applied Biology 127 (1):191-200. doi:DOI 10.1111/j.1744-7348.1995.tb06664.x

McClung CR, Lou P, Hermand V, Kim JA (2016) The Importance of Ambient Temperature to Growth and the Induction of Flowering. Frontiers in Plant Science 7:1266. doi:10.3389/fpls.2016.01266

McDonald GK, Paulsen GM (1997) High temperature effects on photosynthesis and water relations of grain legumes. Plant Soil 196 (1):47-58. doi:Doi 10.1023/A:1004249200050

Melzer S, Lens F, Gennen J, Vanneste S, Rohde A, Beeckman T (2008) Flowering-time genes modulate meristem determinacy and growth form in Arabidopsis thaliana. Nature Genetics 40 (12):1489-1492. doi:10.1038/ng.253

Mena M, Mandel MA, Lerner DR, Yanofsky MF, Schmidt RJ (1995) A characterization of the MADS-box gene family in maize. Plant Journal 8 (6):845-854

Mendel G (1865) Versuche uber pflanzen-hybriden. Verh Naturforsch Ver Brfinn 4:3–47

Meng X, Muszynski MG, Danilevskaya ON (2011) The FT-like ZCN8 gene functions as a floral activator and is involved in photoperiod sensitivity in maize. The Plant Cell 23 (3):942-960. doi:10.1105/tpc.110.081406

Mertes F, ElSharawy A, Sauer S, van Helvoort JMLM, van der Zaag PJ, Franke A, Nilsson M, Lehrach H, Brookes AJ (2011) Targeted enrichment of genomic DNA regions for next-generation sequencing. Briefings in Functional Genomics 10 (6):374-386. doi:10.1093/bfgp/elr033

Michels CA (2002) 6 Epistasis Analysis.

Miko I (2008) Epistasis: Gene interaction and phenotype effects. . Nature Education 1 (1):197

Miller AC, Obholzer ND, Shah AN, Megason SG, Moens CB (2013) RNA-seq–based mapping and candidate identification of mutations from forward genetic screens. Genome Research 23 (4):679-686. doi:10.1101/gr.147322.112

Miller TA, Muslin EH, Dorweiler JE (2008) A maize CONSTANS-like gene, conz1, exhibits distinct diurnal expression patterns in varied photoperiods. Planta 227 (6):1377-1388. doi:10.1007/s00425-008-0709-1

Mishra RK, Gangadhar BH, Nookaraju A, Kumar S, Park SW (2012) Development of EST-derived SSR markers in pea (Pisum sativum) and their potential utility for genetic mapping and transferability. Plant Breeding 131 (1):118-124. doi:10.1111/j.1439-0523.2011.01926.x

Mitchum MG, Yamaguchi S, Hanada A, Kuwahara A, Yoshioka Y, Kato T, Tabata S, Kamiya Y, Sun T-p (2006) Distinct and overlapping roles of two gibberellin 3-oxidases in Arabidopsis development. The Plant Journal 45 (5):804-818. doi:doi:10.1111/j.1365-313X.2005.02642.x

Mitsuda N, Ohme-Takagi M (2009) Functional Analysis of Transcription Factors in Arabidopsis. Plant and Cell Physiology 50 (7):1232-1248. doi:10.1093/pcp/pcp075

Mittler R, Vanderauwera S, Suzuki N, Miller G, Tognetti VB, Vandepoele K, Gollery M, Shulaev V, Van Breusegem F (2011) ROS signaling: the new wave? Trends in Plant Science 16 (6):300-309. doi:10.1016/j.tplants.2011.03.007

Mizuno N, Nitta M, Sato K, Nasuda S (2012) A wheat homologue of PHYTOCLOCK 1 is a candidate gene conferring the early heading phenotype to einkorn wheat. Genes and Genetic Systems 87 (6):357-367

Mohan M, Nair S, Bhagwat A, Krishna TG, Yano M, Bhatia CR, Sasaki T (1997) Genome mapping, molecular markers and marker-assisted selection in crop plants. Molecular Breeding 3 (2):87-103. doi:Doi 10.1023/A:1009651919792

Mohr H (2012) Lectures on photomorphogenesis. Springer Science & Business Media,

Montero-Barrientos M, Hermosa R, Cardoza RE, Gutierrez S, Nicolas C, Monte E (2010) Transgenic expression of the Trichoderma harzianum hsp70 gene increases Arabidopsis resistance to heat and other abiotic stresses. Journal of Plant Physiology 167 (8):659-665. doi:10.1016/j.jplph.2009.11.012

Moreno JE, Tao Y, Chory J, Ballare CL (2009) Ecological modulation of plant defense via phytochrome control of jasmonate sensitivity. Proceedings of the National Academy of Sciences of the United States of America 106 (12):4935-4940. doi:10.1073/pnas.0900701106

Morris SE, Turnbull CGN, Murfet IC, Beveridge CA (2001) Mutational Analysis of Branching in Pea. Evidence That Rms1 and Rms5 Regulate the Same Novel Signal. Plant Physiology 126 (3):1205-1213

Morsy M, Gouthu S, Orchard S, Thorneycroft D, Harper JF, Mittler R, Cushman JC (2008) Charting plant interactomes: possibilities and challenges. Trends in Plant Science 13 (4):183-191. doi:10.1016/j.tplants.2008.01.006

Mount SM (1982) A catalogue of splice junction sequences. Nucleic Acids Research 10 (2):459-472

Mulekar JJ, Bu Q, Chen F, Huq E (2012) Casein kinase II alpha subunits affect multiple developmental and stress-responsive pathways in Arabidopsis. The Plant Journal : for Cell and Molecular Biology 69 (2):343-354. doi:10.1111/j.1365-313X.2011.04794.x

Mulekar JJ, Huq E (2015) Arabidopsis casein kinase 2 alpha4 subunit regulates various developmental pathways in a functionally overlapping manner. Plant Science236:295-303. doi:10.1016/j.plantsci.2015.04.013

Munns R, Tester M (2008) Mechanisms of salinity tolerance. Annual Review of Plant Biology 59:651-681. doi:10.1146/annurev.arplant.59.032607.092911

Murfet I (1971) Flowering in Pisum: reciprocal grafts between known genotypes. Australian Journal of Biological Sciences 24 (4):1089-1102

Murfet I (1975) Flowering in Pisum: Multiple alleles at the Lf locus. Heredity 35 (1):85
Murfet I (1989) Interaction of the det (determinate) mutant with other flowering genes. Pisum Newsletter21:44-47
Murfet I (1992) Garden pea and allies-an update from Hobart. Flowering Newsletter 13:10-20
Murfet R (1993) Developmental mutants, In Peas: Genetics, Molecular Biology and Biotechnology R Casey and DR Davies, eds (Cambridge: CAB International):pp. 165-216
Murphy RL, Klein RR, Morishige DT, Brady JA, Rooney WL, Miller FR, Dugas DV, Klein PE, Mullet JE (2011) Coincident light and clock regulation of pseudoresponse regulator protein 37 (PRR37) controls photoperiodic flowering in sorghum. Proceedings of the National Academy of Sciences of the United States of America 108 (39):16469-16474. doi:10.1073/pnas.1106212108
Nadeem MA, Nawaz MA, Shahid MQ, Doğan Y, Comertpay G, Yıldız M, Hatipoğlu R, Ahmad F, Alsaleh A, Labhane N, Özkan H, Chung G, Baloch FS (2018) DNA molecular markers in plant breeding: current status and recent advancements in genomic selection and genome editing. Biotechnology & Biotechnological Equipment 32 (2):261-285. doi:10.1080/13102818.2017.1400401
Najafi F, Khavari-Nejad R, Rastgar-Jazii F, Sticklen M (2007) Growth and some physiological attributes of pea (Pisum sativum L.) as affected by salinity. Pakistan Journal of Biological Sciences: PJBS 10 (16):2752-2755
Nakajima M, Shimada A, Takashi Y, Kim YC, Park SH, Ueguchi-Tanaka M, Suzuki H, Katoh E, Iuchi S, Kobayashi M, Maeda T, Matsuoka M, Yamaguchi I (2006) Identification and characterization of Arabidopsis gibberellin receptors. The Plant Journal : for Cell and Molecular Biology 46 (5):880-889. doi:10.1111/j.1365-313X.2006.02748.x
Narvaez-Vasquez J, Ryan CA (2004) The cellular localization of prosystemin: a functional role for phloem parenchyma in systemic wound signaling. Planta 218 (3):360-369. doi:10.1007/s00425-003-1115-3
Neff MM, Turk E, Kalishman M (2002) Web-based primer design for single nucleotide polymorphism analysis. Trends in Genetics 18 (12):613-615. doi:Pii S0168-9525(02)02820-2

Doi 10.1016/S0168-9525(02)02820-2
Nemet J, Jelicic B, Rubelj I, Sopta M (2014) The two faces of Cdk8, a positive/negative regulator of transcription. Biochimie 97:22-27. doi:10.1016/j.biochi.2013.10.004
Nemhauser J, Chory J (2002) Photomorphogenesis. The Arabidopsis Book / American Society of Plant Biologists 1:e0054. doi:10.1199/tab.0054
Neumann P, Nouzova M, Macas J (2001) Molecular and cytogenetic analysis of repetitive DNA in pea (Pisum sativum L.). Genome 44 (4):716-728. doi:DOI 10.1139/gen-44-4-716
Ng S, Giraud E, Duncan O, Law SR, Wang Y, Xu L, Narsai R, Carrie C, Walker H, Day DA, Blanco NE, Strand A, Whelan J, Ivanova A (2013) Cyclin-dependent Kinase E1 (CDKE1) Provides a Cellular Switch in Plants between Growth and Stress Responses. Journal of Biological Chemistry 288 (5):3449-3459. doi:10.1074/jbc.M112.416727
Nilsen TW (2003) The spliceosome: the most complex macromolecular machine in the cell? BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology 25 (12):1147-1149. doi:10.1002/bies.10394
Nodine MD, Yadegari R, Tax FE (2007) RPK1 and TOAD2 are two receptor-like kinases redundantly required for arabidopsis embryonic pattern formation. Developmental cell 12 (6):943-956. doi:10.1016/j.devcel.2007.04.003
O'Malley RC, Barragan CC, Ecker JR (2015) A User's Guide to the Arabidopsis T-DNA Insertional Mutant Collections. Methods in Molecular Biology (Clifton, NJ) 1284:323-342. doi:10.1007/978-1-4939-2444-8_16
Oaks A (1992) A re-evaluation of nitrogen assimilation in roots. Bioscience 42 (2):103-111
Obholzer N, Swinburne IA, Schwab E, Nechiporuk AV, Nicolson T, Megason SG (2012) Rapid positional cloning of zebrafish mutations by linkage and homozygosity mapping using whole-genome sequencing. Development 139 (22):4280-4290. doi:10.1242/dev.083931
Oh S, Park S, Han KH (2003) Transcriptional regulation of secondary growth in Arabidopsis thaliana. Journal of Experimental Botany 54 (393):2709-2722. doi:10.1093/jxb/erg304
Ohama N, Sato H, Shinozaki K, Yamaguchi-Shinozaki K (2017) Transcriptional Regulatory Network of Plant Heat Stress Response. Trends in Plant Science 22 (1):53-65. doi:10.1016/j.tplants.2016.08.015
Okamuro JK, denBoer BGW, LotysPrass C, Szeto W, Jofuku KD (1996) Flowers into shoots: Photo and hormonal control of a meristem identity switch in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America 93 (24):13831-13836. doi:DOI 10.1073/pnas.93.24.13831

Okamuro JK, Szeto W, LotysPrass C, Jofuku KD (1997) Photo and hormonal control of meristem identity in the Arabidopsis flower mutants apetala2 and apetala1. The Plant Cell 9 (1):37-47

Olson SN, Ritter K, Rooney W, Kemanian A, McCarl BA, Zhang Y, Hall S, Packer D, Mullet J (2012) High biomass yield energy sorghum: Developing a genetic model for C4 grass bioenergy crops. Biofuels, Bioproducts and Biorefining 6 (6):640-655. doi:10.1002/bbb.1357

Ordidge M, Chiurugwi T, Tooke F, Battey NH (2005) LEAFY, TERMINAL FLOWER1 and AGAMOUS are functionally conserved but do not regulate terminal flowering and floral determinacy in Impatiens balsamina. The Plant Journal 44 (6):985-1000

Osnato M, Castillejo C, Matias-Hernandez L, Pelaz S (2012) TEMPRANILLO genes link photoperiod and gibberellin pathways to control flowering in Arabidopsis. Nature communications 3:808. doi:10.1038/ncomms1810

Ozga JA, Reinecke DM (2003) Hormonal interactions in fruit development. Journal of Journal of Plant Growth Regulationation 22 (1):73-81. doi:10.1007/s00344-003-0024-9

Ozturk M, Szaniawski RK (1981) Root Temperature Stress and Proline Content in Leaves and Roots of 2 Ecologically Different Plant-Species. Zeitschrift Fur Pflanzenphysiologie 102 (4):375-377

P J, Phillips CaL, Croker SJ, García-Lepe R, Lewis MJ, Hedden P (1999) Modification of gibberellin production and plant development in Arabidopsis by sense and antisense expression of gibberellin 20-oxidase genes. The Plant Journal 17 (5):547-556. doi:doi:10.1046/j.1365-313X.1999.00410.x

Pacifici E, Polverari L, Sabatini S (2015) Plant hormone cross-talk: the pivot of root growth. Journal of Experimental Botany 66 (4):1113-1121. doi:10.1093/jxb/eru534

Pagliarini MS (2000) Meiotic behavior of economically important plant species: the relationship between fertility and male sterility. Genetics and Molecular Biology 23 (4):997-1002. doi:Doi 10.1590/S1415-47572000000400045

Parcy F, Bomblies K, Weigel D (2002) Interaction of LEAFY, AGAMOUS and TERMINAL FLOWER1 in maintaining floral meristem identity in Arabidopsis. Development 129 (10):2519-2527

Park S-W, Do H-J, Ha WT, Han M-H, Park K-H, Song H, Kim N-H, Kim J-H (2014) Transcriptional activation of OCT4 by the ETS transcription factor PEA3 in NCCIT human embryonic carcinoma cells. FEBS Letters 588 (17):3129-3136. doi:https://doi.org/10.1016/j.febslet.2014.06.052

Pasrija R, Thakur JK (2012) Analysis of differential expression of Mediator subunit genes in Arabidopsis. Plant Signaling & Behavior 7 (12):1676-1686. doi:10.4161/psb.22438

Pate J (1977) The pea as a crop plant. In The Physiology of the Garden Pea Eds JF Sutcliffe and JS Pate:469-484 Academic press, London

Pate J (1980) Transport and partitioning of nitrogenous solutes. Annual Review of Plant Physiology 31 (1):313-340

Penacortes H, Fisahn J, Willmitzer L (1995) Signals Involved in Wound-Induced Proteinase-Inhibitor-Ii Gene-Expression in Tomato and Potato Plants. Proceedings of the National Academy of Sciences of the United States of America 92 (10):4106-4113. doi:DOI 10.1073/pnas.92.10.4106

Pertea M, Mount SM, Salzberg SL (2007) A computational survey of candidate exonic splicing enhancer motifs in the model plant Arabidopsis thaliana. BMC Bioinformatics 8 (1):159. doi:10.1186/1471-2105-8-159

Peters JL, Cnudde F, Gerats T (2003) Forward genetics and map-based cloning approaches. Trends in Plant Science 8 (10):484-491. doi:https://doi.org/10.1016/j.tplants.2003.09.002

Pflieger S, Lefebvre V, Causse M (2001) The candidate gene approach in plant genetics: a review. Molecular Breeding 7 (4):275-291. doi:10.1023/a:1011605013259

Pien S, Grossniklaus U (2007) Polycomb group and trithorax group proteins in Arabidopsis. BBA - Gene Structure and Expression 1769 (5-6):375-382. doi:10.1016/j.bbaexp.2007.01.010

Pin PA, Nilsson O (2012) The multifaceted roles of FLOWERING LOCUS T in plant development. Plant, Cell & Environment 35 (10):1742-1755. doi:10.1111/j.1365-3040.2012.02558.x

Plackett ARG, Powers SJ, Fernandez-Garcia N, Urbanova T, Takebayashi Y, Seo M, Jikumaru Y, Benlloch R, Nilsson O, Ruiz-Rivero O, Phillips AL, Wilson ZA, Thomas SG, Hedden P (2012) Analysis of the developmental roles of the Arabidopsis gibberellin 20-oxidases demonstrates that GA20ox1, -2, and -3 are the dominant paralogs. The Plant Cell 24 (3):941-960. doi:10.1105/tpc.111.095109

Platten JD, Foo E, Elliott RC, Hecht V, Reid JB, Weller JL (2005) Cryptochrome 1 contributes to blue-light sensing in pea. Plant Physiology 139 (3):1472-1482. doi:DOI 10.1104/pp.105.067462

Pouteau S, Nicholls D, Tooke F, Coen E, Battey N (1997) The induction and maintenance of flowering in Impatiens. Development 124 (17):3343-3351

Prasad PVV, Boote KJ, Allen LH (2006) Adverse high temperature effects on pollen viability, seed-set, seed yield and harvest index of grain-sorghum [Sorghum bicolor (L.) Moench] are more severe at elevated carbon

dioxide due to higher tissue temperatures. Agricultural and Forest Meteorology 139 (3-4):237-251. doi:10.1016/j.agrformet.2006.07.003

Qin J, Ma X, Yi Z, Tang Z, Meng Y (2016) A transcriptome-wide study on the microRNA- and the Argonaute 1-enriched small RNA-mediated regulatory networks involved in plant leaf senescence. Plant Biology 18 (2):197-205. doi:10.1111/plb.12373

Qiu QS, Guo Y, Dietrich MA, Schumaker KS, Zhu JK (2002) Regulation of SOS1, a plasma membrane Na+/H+ exchanger in Arabidopsis thaliana, by SOS2 and SOS3. Proceedings of the National Academy of Sciences of the United States of America 99 (12):8436-8441. doi:10.1073/pnas.122224699

Quint M, Delker C, Franklin KA, Wigge PA, Halliday KJ, van Zanten M (2016) Molecular and genetic control of plant thermomorphogenesis. Nature Plants 2 (1). doi:Artn 1519010.1038/Nplants.2015.190

Rameau C, Bertheloot J, Leduc N, Andrieu B, Foucher F, Sakr S (2015) Multiple pathways regulate shoot branching. Frontiers in Plant Science 5:741

Randriamanana TR, Nissinen K, Moilanen J, Nybakken L, Julkunen-Tiitto R (2015) Long-term UV-B and temperature enhancements suggest that females of Salix myrsinifolia plants are more tolerant to UV-B than males. Environmental and Experimental Botany 109:296-305. doi:https://doi.org/10.1016/j.envexpbot.2014.06.007

Ratcliffe OJ, Amaya I, Vincent CA, Rothstein S, Carpenter R, Coen ES, Bradley DJ (1998) A common mechanism controls the life cycle and architecture of plants. Development 125 (9):1609-1615

Ream TS, Woods DP, Amasino RM (2012) The molecular basis of vernalization in different plant groups. Cold Spring Harbor Symposia on Quantitative Biology 77:105-115. doi:10.1101/sqb.2013.77.014449

Reid JB, Murfet IC, Singer SR, Weller JL, Taylor SA Physiological-genetics of flowering inPisum. In: Seminars in Cell & Developmental Biology, 1996. vol 3. Elsevier, pp 455-463

Reid JB RJ (1988) Internode length in Pisum. A new gene, Iv, conferring an increased response to gibberellin A. Physiologia Plantarum 72:595-604

Reinecke DM (1999) 4-chloroindole-3-acetic acid and plant growth. Journal of Plant Growth Regulation 27 (1):3-13. doi:Doi 10.1023/A:1006191917753

Richards RA, Hunt JR, Kirkegaard JA, Passioura JB (2014) Yield improvement and adaptation of wheat to water-limited environments in Australia-a case study. Crop Pasture Sci 65 (7):676-689. doi:10.1071/Cp13426

Ridge S, Sussmilch FC, Hecht V, Vander Schoor JK, Lee R, Aubert G, Burstin J, Macknight RC, Weller JL (2016) Identification of LATE BLOOMER2 as a CYCLING DOF FACTOR Homolog Reveals Conserved and Divergent Features of the Flowering Response to Photoperiod in Pea. The Plant Cell 28 (10):2545-2559. doi:10.1105/tpc.15.01011

Riemann M, Bouyer D, Hisada A, Muller A, Yatou O, Weiler EW, Takano M, Furuya M, Nick P (2009) Phytochrome A requires jasmonate for photodestruction. Planta 229 (5):1035-1045. doi:10.1007/s00425-009-0891-9

Riemann M, Riemann M, Takano M (2008) Rice JASMONATE RESISTANT 1 is involved in phytochrome and jasmonate signalling. Plant Cell and Environment 31 (6):783-792. doi:10.1111/j.1365-3040.2008.01790.x

Rieu I, Ruiz-Rivero O, Fernandez-Garcia N, Griffiths J, Powers SJ, Gong F, Linhartova T, Eriksson S, Nilsson O, Thomas SG, Phillips AL, Hedden P (2008) The gibberellin biosynthetic genes AtGA20ox1 and AtGA20ox2 act, partially redundantly, to promote growth and development throughout the Arabidopsis life cycle. The Plant Journal 53 (3):488-504. doi:doi:10.1111/j.1365-313X.2007.03356.x

Risch N (1992) Genetic-Linkage - Interpreting Lod Scores. Science 255 (5046):803-804. doi:DOI 10.1126/science.1536004

Rizhsky L, Liang HJ, Shuman J, Shulaev V, Davletova S, Mittler R (2004) When Defense pathways collide. The response of Arabidopsis to a combination of drought and heat stress. Plant Physiology 134 (4):1683-1696. doi:10.1104/pp.103.033431

Rizzini L, Favory J-J, Cloix C, Faggionato D, O'Hara A, Kaiserli E, Baumeister R, Schäfer E, Nagy F, Jenkins GI (2011) Perception of UV-B by the Arabidopsis UVR8 protein. Science 332 (6025):103-106

Robin AHK, Matthew C, Uddin MJ, Bayazid KN (2016) Salinity-induced reduction in root surface area and changes in major root and shoot traits at the phytomer level in wheat. Journal of Experimental Botany 67 (12):3719-3729. doi:10.1093/jxb/erw064

Robles P, Pelaz S (2005) Flower and fruit development in Arabidopsis thaliana.The International Journal of Developmental Biology 49 (5-6):633-643. doi:10.1387/ijdb.052020pr

Robson F, Okamoto H, Patrick E, Harris SR, Wasternack C, Brearley C, Turner JG (2010) Jasmonate and Phytochrome A Signaling in Arabidopsis Wound and Shade Responses Are Integrated through JAZ1 Stability. The Plant Cell 22 (4):1143-1160. doi:10.1105/tpc.109.067728

Robson TM, Klem K, Urban O, Jansen MAK (2015) Re-interpreting plant morphological responses to UV-B radiation. Plant Cell and Environment 38 (5):856-866. doi:10.1111/pce.12374

Robson TM, Pancotto VA, Flint SD, Ballaré CL, Sala OE, Scopel AL, Caldwell MM (2003) Six years of solar UV‐B manipulations affect growth of Sphagnum and vascular plants in a Tierra del Fuego peatland. New Phytologist 160 (2):379-389

Roca X, Sachidanandam R, Krainer AR (2003) Intrinsic differences between authentic and cryptic 5 ' splice sites. Nucleic Acids Research 31 (21):6321-6333. doi:10.1093/nar/gkg830

Rochange S (2010) Strigolactones and their role in arbuscular mycorrhizal symbiosis. In: Arbuscular Mycorrhizas: Physiology and Function. Springer, pp 73-90

Rogan PK, Faux BM, Schneider TD (1998) Information analysis of human splice site mutations. Human Mutation 12 (3):153-171

Rooney WL, Aydin S (1999) Genetic control of a photoperiod-sensitive response in Sorghum bicolor (L.) Moench. Crop Science 39 (2):397-400

Rooney WL, Blumenthal J, Bean B, Mullet JE (2007) Designing sorghum as a dedicated bioenergy feedstock. Biofuels, Bioproducts and Biorefining 1 (2):147-157. doi:10.1002/bbb.15

Ross JJ, O'Neill DP, Rathbone DA (2003) Auxin-gibberellin interactions in pea: Integrating the old with the new. Journal of Plant Growth Regulation 22 (1):99-108. doi:10.1007/s00344-003-0021-z

Ross JJ, O'Neill DP, Smith JJ, Kerckhoffs LHJ, Elliott RC (2000) Evidence that auxin promotes gibberellin A(1) biosynthesis in pea. Plant Journal 21 (6):547-552. doi:DOI 10.1046/j.1365-313x.2000.00702.x

Ross JJ, Reid jB (1992) Ontogenetic and environmental effects on GA1 levels and the implications for the control of internode length. In: Karssen CM, van Loon LC, Vreugdenhil D (eds) Progress in Journal of Plant Growth Regulationation: Proceedings of the 14th International Conference on Plant Growth Substances, Amsterdam, 21–26 July, 1991. Springer Netherlands, Dordrecht, pp 180-187. doi:10.1007/978-94-011-2458-4_19

Ross JJ, Reid JB, Dungey HS (1992a) Ontogenetic variation in levels of gibberellin A1 in Pisum : Implications for the control of stem elongation. Planta 186 (2):166-171. doi:10.1007/BF00196245

Ross JJ, Reid JB, Dungey HS (1992b) Ontogenetic variation in levels of gibberellin A 1 in Pisum. Planta 186 (2):166-171

Rubenach AJS, Hecht VF, Vander Schoor JK, Liew LC, Aubert G, Burstin J, Weller JL (2017) Early Flowering 3 redundancy fine-tunes photoperiod sensitivity. Plant Physiology. doi:10.1104/pp.16.01738

Ruelens P, de Maagd RA, Proost S, Theissen G, Geuten K, Kaufmann K (2013) FLOWERING LOCUS C in monocots and the tandem origin of angiosperm-specific MADS-box genes. Nature Communications 4:2280. doi:10.1038/ncomms3280

Rugnone ML, Soverna AF, Sanchez SE, Schlaen RG, Hernando CE, Seymour DK, Mancini E, Chernomoretz A, Weigel D, Maac P, Yanovsky MJ (2013) LNK genes integrate light and clock signaling networks at the core of the Arabidopsis oscillator. Proceedings of the National Academy of Sciences of the United States of America 110 (29):12120-12125. doi:DOI 10.1073/pnas.1302170110

Sadras VO, Lake L, Chenu K, McMurray LS, Leonforte A (2012) Water and thermal regimes for field pea in Australia and their implications for breeding. Crop Pasture Sci 63 (1):33-44. doi:10.1071/Cp11321

Sage J (2004) Cyclin C makes an entry into the cell cycle. Developmental cell 6 (5):607-608. doi:Doi 10.1016/S1534-5807(04)00137-6

Sakamoto H, Maruyama K, Sakuma Y, Meshi T, Iwabuchi M, Shinozaki K, Yamaguchi-Shinozaki K (2004) Arabidopsis Cys2/His2-type zinc-finger proteins function as transcription repressors under drought, cold, and high-salinity stress conditions. Plant Physiology 136 (1):2734-2746. doi:10.1104/pp.104.046599

Samach A, Onouchi H, Gold SE, Ditta GS, Schwarz-Sommer Z, Yanofsky MF, Coupland G (2000) Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis. Science 288 (5471):1613-1616. doi:DOI 10.1126/science.288.5471.1613

Samanta S, Thakur JK (2015) Importance of Mediator complex in the regulation and integration of diverse signaling pathways in plants. Frontiers in Plant Science 6:757. doi:10.3389/fpls.2015.00757

Sanchez P, Nehlin L, Greb T (2012) From thin to thick: major transitions during stem development. Trends in Plant Science 17 (2):113-121. doi:https://doi.org/10.1016/j.tplants.2011.11.004

Sanders A, Collier R, Trethewy A, Gould G, Sieker R, Tegeder M (2009) AAP1 regulates import of amino acids into developing Arabidopsis embryos. Plant Journal 59 (4):540-552. doi:10.1111/j.1365-313X.2009.03890.x

Sanders I, Sukharnikov L, Najar FZ, Roe BA (2011) Medicago. In: Wild Crop Relatives: Genomic and Breeding Resources. Springer, pp 207-222

Sato S, Nakamura Y, Asamizu E, Isobe S, Tabata S (2007) Genome sequencing and genome resources in model legumes. Plant Physiology 144 (2):588-593. doi:10.1104/pp.107.097493

Savatin DV, Gramegna G, Modesti V, Cervone F (2014) Wounding in the plant tissue: the defense of a dangerous passage. Frontiers in Plant Science 5 (470). doi:10.3389/fpls.2014.00470

Sawa M, Nusinow DA, Kay SA, Imaizumi T (2007) FKF1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis. Science 318 (5848):261-265. doi:DOI 10.1126/science.1146994

Schilmiller AL, Howe GA (2005) Systemic signaling in the wound response. Current Opinion in Plant Biology 8 (4):369-377. doi:10.1016/j.pbi.2005.05.008

Schmidt R, Stransky H, Koch W (2007) The amino acid permease AAP8 is important for early seed development in Arabidopsis thaliana. Planta 226 (4):805-813. doi:10.1007/s00425-007-0527-x

Schmitz RJ, Hong L, Michaels S, Amasino RM (2005) FRIGIDA-ESSENTIAL 1 interacts genetically with FRIGIDA and FRIGIDA-LIKE 1 to promote the winter-annual habit of Arabidopsis thaliana. Development 132 (24):5471-5478. doi:10.1242/dev.02170

Schneeberger K (2014) Using next-generation sequencing to isolate mutant genes from forward genetic screens. Nature Reviews Genetics 15:662. doi:10.1038/nrg3745

Schomburg FM, Bizzell CM, Lee DJ, Zeevaart JAD, Amasino RM (2003) Overexpression of a novel class of gibberellin 2-oxidases decreases gibberellin levels and creates dwarf plants. The Plant Cell 15 (1):151-163. doi:10.1105/tpc.005975

Schomburg FM, Patton DA, Meinke DW, Amasino RM (2001) FPA, a gene involved in floral induction in Arabidopsis, encodes a protein containing RNA-recognition motifs. The Plant Cell 13 (6):1427-1436

Schroeder HE, Schotz AH, Wardley-Richardson T, Spencer D, Higgins T (1993) Transformation and Regeneration of Two Cultivars of Pea (Pisum sativum L.). Plant Physiology 101 (3):751-757. doi:10.1104/pp.101.3.751

Sehgal A, Sita K, Kumar J, Kumar S, Singh S, Siddique KHM, Nayyar H (2017) Effects of Drought, Heat and Their Interaction on the Growth, Yield and Photosynthetic Function of Lentil (Lens culinaris Medikus) Genotypes Varying in Heat and Drought Sensitivity. Frontiers in Plant Science 8:1776-1776. doi:10.3389/fpls.2017.01776

Sehr EM, Agusti J, Lehner R, Farmer EE, Schwarz M, Greb T (2010) Analysis of secondary growth in the Arabidopsis shoot reveals a positive role of jasmonate signalling in cambium formation. Plant Journal 63 (5):811-822. doi:10.1111/j.1365-313X.2010.04283.x

Semenov MA, Halford NG (2009) Identifying target traits and molecular mechanisms for wheat breeding under a changing climate. Journal of Experimental Botany 60 (10):2791-2804. doi:10.1093/jxb/erp164

Sewelam N, Oshima Y, Mitsuda N, OHME‐TAKAGI M (2014) A step towards understanding plant responses to multiple environmental stresses: a genome‐wide study. Plant, Cell & Environment 37 (9):2024-2035

Shani E, Weinstain R, Zhang Y, Castillejo C, Kaiserli E, Chory J, Tsien RY, Estelle M (2013) Gibberellins accumulate in the elongating endodermal cells of Arabidopsis root. Proceedings of the National Academy of Sciences 110 (12):4834-4839

Shannon S, Meeks-Wagner DR (1993) Genetic Interactions That Regulate Inflorescence Development in Arabidopsis. The Plant Cell 5 (6):639-655

Shen WH, Xu L (2009) Chromatin Remodeling in Stem Cell Maintenance in Arabidopsis thaliana. Molecular Plant 2 (4):600-609. doi:10.1093/mp/ssp022

Sheth BP, Thaker VS (2014) Plant systems biology: insights, advances and challenges. Planta 240 (1):33-54. doi:10.1007/s00425-014-2059-5

Shim JS, Imaizumi T (2015) Circadian clock and photoperiodic response in Arabidopsis: from seasonal flowering to redox homeostasis. Biochemistry 54 (2):157-170. doi:10.1021/bi500922q

Shim JS, Kubota A, Imaizumi T (2017) Circadian Clock and Photoperiodic Flowering in Arabidopsis: CONSTANS Is a Hub for Signal Integration. Plant Physiology 173 (1):5-15. doi:10.1104/pp.16.01327

Shimada S, Ogawa T, Kitagawa S, Suzuki T, Ikari C, Shitsukawa N, Abe T, Kawahigashi H, Kikuchi R, Handa H, Murai K (2009) A genetic network of flowering-time genes in wheat leaves, in which an APETALA1/FRUITFULL-like gene, VRN1, is upstream of FLOWERING LOCUS T. The Plant Journal : for Cell and Molecular Biology 58 (4):668-681. doi:10.1111/j.1365-313X.2009.03806.x

Shivanna KR, Rangaswamy NS (1992) Tests for Pollen Viability. In: Pollen Biology: A Laboratory Manual. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 33-37. doi:10.1007/978-3-642-77306-8_5

Sibomana I, Aguyoh J, Opiyo A (2013) Water stress affects growth and yield of container grown tomato (Lycopersicon esculentum Mill) plants. Global Journal of Bio-Science and BioTechnology 2 (4):461-466

Sikorski TW, Buratowski S (2009) The basal initiation machinery: beyond the general transcription factors. Current Opinion in Cell Biology 21 (3):344-351. doi:10.1016/j.ceb.2009.03.006

Simons JL, Napoli CA, Janssen BJ, Plummer KM, Snowden KC (2007) Analysis of the DECREASED APICAL DOMINANCE genes of petunia in the control of axillary branching. Plant Physiology 143 (2):697-706. doi:10.1104/pp.106.087957

Simpson GG, Dijkwel PP, Quesada V, Henderson I, Dean C (2003) FY is an RNA 3 ' end-processing factor that interacts with FCA to control the Arabidopsis floral transition. Cell 113 (6):777-787. doi:Doi 10.1016/S0092-8674(03)00425-2

Sindhu A, Ramsay L, Sanderson L-A, Stonehouse R, Li R, Condie J, Shunmugam ASK, Liu Y, Jha AB, Diapari M, Burstin J, Aubert G, Tar'an B, Bett KE, Warkentin TD, Sharpe AG (2014) Gene-based SNP discovery and genetic mapping in pea. TAG Theoretical and applied genetics Theoretische und angewandte Genetik 127 (10):2225-2241. doi:10.1007/s00122-014-2375-y

Singer SR, Hsiung LP, Huber SC (1990) Determinate (det) mutant of Pisum sativum (Leguminosae: Papilionoideae) exhibits an indeterminate growth pattern. American Journal of Botany 77 (10):1330-1335

Smillie RM, Critchley C, Bain JM, Nott R (1978) Effect of growth temperature on chloroplast structure and activity in barley. Plant Physiology 62 (2):191-196

Smýkal P, Aubert G, Burstin J, Coyne CJ, Ellis NTH, Flavell AJ, Ford R, Hýbl M, Macas J, Neumann P, McPhee KE, Redden RJ, Rubiales D, Weller JL, Warkentin TD (2012) Pea (Pisum sativum L.) in the Genomic Era. Agronomy 2 (4):74-115. doi:10.3390/agronomy2020074

Snowden KC, Simkin AJ, Janssen BJ, Templeton KR, Loucas HM, Simons JL, Karunairetnam S, Gleave AP, Clark DG, Klee HJ (2005) The Decreased apical dominance 1/petunia hybrida carotenoid cleavage dioxygenase8 gene affects branch production and plays a role in leaf senescence, root growth, and flower development. The Plant Cell 17 (3):746-759. doi:10.1105/tpc104.027714

Somers J, Pöyry T, Willis AE (2013) A perspective on mammalian upstream open reading frame function. The International Journal of Biochemistry & Cell Biology 45 (8):1690-1700. doi:https://doi.org/10.1016/j.biocel.2013.04.020

Song G, Jia M, Chen K, Kong X, Khattak B, Xie C, Li A, Mao L (2016) CRISPR/Cas9: A powerful tool for crop genome editing. The Crop Journal 4 (2):75-82. doi:https://doi.org/10.1016/j.cj.2015.12.002

Song YH, Ito S, Imaizumi T (2013) Flowering time regulation: photoperiod- and temperature-sensing in leaves. Trends in Plant Science 18 (10):575-583. doi:10.1016/j.tplants.2013.05.003

Song YH, Lee I, Lee SY, Imaizumi T, Hong JC (2012a) CONSTANS and ASYMMETRIC LEAVES 1 complex is involved in the induction of FLOWERING LOCUS T in photoperiodic flowering in Arabidopsis. The Plant Journal 69 (2):332-342. doi:10.1111/j.1365-313X.2011.04793.x

Song YH, Shim JS, Kinmonth-Schultz HA, Imaizumi T (2015) Photoperiodic Flowering: Time Measurement Mechanisms in Leaves. Annual Review of Plant Biology, Vol 66 66:441-464. doi:10.1146/annurev-arplant-043014-115555

Song YH, Smith RW, To BJ, Millar AJ, Imaizumi T (2012b) FKF1 conveys timing information for CONSTANS stabilization in photoperiodic flowering. Science 336 (6084):1045-1049. doi:10.1126/science.1219644

Soni P, Nutan KK, Soda N, Nongpiur RC, Roy S, Singla-Pareek SL, Pareek A (2015) Towards understanding abiotic stress signaling in plants: convergence of genomic, transcriptomic, proteomic, and metabolomic approaches. In: Elucidation of Abiotic Stress Signaling in Plants. Springer, pp 3-40

Sorefan K, Booker J, Haurogne K, Goussot M, Bainbridge K, Foo E, Chatfield S, Ward S, Beveridge C, Rameau C, Leyser O (2003) MAX4 and RMS1 are orthologous dioxygenase-like genes that regulate shoot branching in Arabidopsis and pea. Genes & Development17 (12):1469-1474. doi:10.1101/gad.256603

Sotelo-Silveira M, Marsch-Martinez N, de Folter S (2014) Unraveling the signal scenario of fruit set. Planta 239 (6):1147-1158. doi:10.1007/s00425-014-2057-7

Springer NM, Danilevskaya ON, Hermon P, Helentjaris TG, Phillips RL, Kaeppler HF, Kaeppler SM (2002) Sequence Relationships, Conserved Domains, and Expression Patterns for Maize Homologs of the Polycomb Group Genes E(z), esc, and E(Pc). Plant Physiology 128 (4):1332-1345. doi:10.1104/pp.010742

Springer NM, Napoli CA, Selinger DA, Pandey R, Cone KC, Chandler VL, Kaeppler HF, Kaeppler SM (2003) Comparative analysis of SET domain proteins in maize and Arabidopsis reveals multiple duplications preceding the divergence of monocots and dicots. Plant Physiology 132 (2):907-925. doi:10.1104/pp.102.013722

Srivastava R, Zhang Y, Xiong X, Zhang X, Pan X, Dong XC, Liangpunsakul S, Janga SC (2016) Prediction and Validation of Transcription Factors Modulating the Expression of Sestrin3 Gene Using an Integrated Computational and Experimental Approach. PLOS One 11 (7):e0160228. doi:10.1371/journal.pone.0160228

Stam P (1993) Construction of integrated genetic linkage maps by means of a new computer package: Join Map. The Plant Journal 3 (5):739-744

Stanfield B, Ormrod D, Fletcher H (1966) Response of peas to environment: II. Effects of temperature in controlled-environment cabinets. Canadian Journal of Plant Science46 (2):195-203

Suarez-Lopez P, Wheatley K, Robson F, Onouchi H, Valverde F, Coupland G (2001) CONSTANS mediates between the circadian clock and the control of flowering in Arabidopsis. Nature 410 (6832):1116-1120

Subbarao GV, Ito O, Berry WL, Wheeler RM (2003) Sodium - A functional plant nutrient. Crit Rev Plant Science22 (5):391-416. doi:10.1080/07352680390243495

Subbarao GV, Wheeler RM, Stutte GW, Levine LH (1999) How far can sodium substitute for potassium in red beet? Journal of Plant Nutrition 22 (11):1745-1761. doi:Doi 10.1080/01904169909365751

Sullivan JA, Gray JC (2000) The Pea light-independent photomorphogenesis1 Mutant Results from Partial Duplication of COP1 Generating an Internal Promoter and Producing Two Distinct Transcripts. The Plant Cell 12 (10):1927-1938

Summerer D (2009) Enabling technologies of genomic-scale sequence enrichment for targeted high-throughput sequencing. Genomics 94 (6):363-368. doi:https://doi.org/10.1016/j.ygeno.2009.08.012

Summerfield RJ, Roberts EH (1988) Photo-thermal regulation of flowering in pea, lentil, faba bean and chickpea. In: Summerfield RJ (ed) World crops: Cool season food legumes: A global perspective of the problems and prospects for crop improvement in pea, lentil, faba bean and chickpea. Springer Netherlands, Dordrecht, pp 911-922. doi:10.1007/978-94-009-2764-3_72

Sun X, Yang T, Hao J, Zhang X, Ford R, Jiang J, Wang F, Guan J, Zong X (2014) SSR genetic linkage map construction of pea (Pisum sativum L.) based on Chinese native varieties. The Crop Journal 2 (2):170-174. doi:https://doi.org/10.1016/j.cj.2014.03.004

Sunarpi, Horie T, Motoda J, Kubo M, Yang H, Yoda K, Horie R, Chan WY, Leung HY, Hattori K, Konomi M, Osumi M, Yamagami M, Schroeder JI, Uozumi N (2005) Enhanced salt tolerance mediated by AtHKT1 transporter-induced Na unloading from xylem vessels to xylem parenchyma cells. The Plant Journal : for Cell and Molecular Biology 44 (6):928-938. doi:10.1111/j.1365-313X.2005.02595.x

Surabhi G-K, Reddy KR, Singh SK (2009) Photosynthesis, fluorescence, shoot biomass and seed weight responses of three cowpea (Vigna unguiculata (L.) Walp.) cultivars with contrasting sensitivity to UV-B radiation. Environmental and Experimental Botany66 (2):160-171. doi:https://doi.org/10.1016/j.envexpbot.2009.02.004

Susete AC, Grégoire A, Sébastien C, Corinne C, Anne‐Lise B, Françoise J, Anthony K, Chantal M, Karen B, Jonathan K, Corinne S, Sandra M, Pascal G, Patrick W, Jérôme G, Judith B (2015) Full‐length de novo assembly of RNA‐seq data in pea (Pisum sativum L.) provides a gene expression atlas and gives insights into root nodulation in this species. The Plant Journal 84 (1):1-19. doi:doi:10.1111/tpj.12967

Susila H, Jin S, Ahn JH (2016) Light Intensity and Floral Transition: Chloroplast Says "Time to Flower!". Molecular Plant 9 (12):1551-1553. doi:10.1016/j.molp.2016.10.013

Sussmilch F (2014) Genetic control of inflorescence development in pea (Doctoral thesis). Retrieved from University of Tasmania Morris Miller Library Database

Sussmilch FC, Berbel A, Hecht V, Vander Schoor JK, Ferrándiz C, Madueño F, Weller JL (2015) Pea VEGETATIVE2 Is an FD Homolog That Is Essential for Flowering and Compound Inflorescence Development. The Plant Cell 27 (4):1046-1060. doi:10.1105/tpc.115.136150

Suzuki N, Mittler R (2012) Reactive oxygen species-dependent wound responses in animals and plants. Free Radical Biology and Medicine 53 (12):2269-2276. doi:10.1016/j.freeradbiomed.2012.10.538

Švábová L, Griga M (2008) The effect of cocultivation treatments on transformation efficiency in pea (Pisum sativum L.). Plant Cell, Tissue and Organ Culture 95 (3):293-304. doi:10.1007/s11240-008-9443-4

Swiecicki W (1987) determinate growth (det) in Pisum: A new mutant gene on chromosome 7. Pisum Newsletter 19:72-73

Swiezewski S, Liu F, Magusin A, Dean C (2009) Cold-induced silencing by long antisense transcripts of an Arabidopsis Polycomb target. Nature 462 (7274):799-802. doi:10.1038/nature08618

Swofford DL (2001) Paup*: Phylogenetic analysis using parsimony (and other methods) 4.0. B5.

Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, Simonovic M, Roth A, Santos A, Tsafou KP, Kuhn M, Bork P, Jensen LJ, von Mering C (2015) STRING v10: protein–protein interaction networks, integrated over the tree of life. Nucleic Acids Research 43 (Database issue):D447-D452. doi:10.1093/nar/gku1003

Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P (2016) The STRING database in 2017: quality-controlled protein–protein association networks, made broadly accessible. Nucleic Acids Research:gkw937

Taatjes DJ, Marr MT, Tjian R (2004) Regulatory diversity among metazoan co-activator complexes. Nature Reviews Molecular Cell Biology 5 (5):403-410. doi:10.1038/nrm1369

Tadege M, Wen J, He J, Tu H, Kwak Y, Eschstruth A, Cayrel A, Endre G, Zhao PX, Chabaud M (2008) Large‐scale insertional mutagenesis using the Tnt1 retrotransposon in the model legume Medicago truncatula. The Plant Journal 54 (2):335-347

Tai YS (2008) Interactome of signaling networks in wheat: the protein-protein interaction between TaRAR1 and TaSGT1. Molecular Biology Reports35 (3):337-343. doi:10.1007/s11033-007-9091-5

Takada S, Jurgens G (2007) Transcriptional regulation of epidermal cell fate in the Arabidopsis embryo. Development 134 (6):1141-1150. doi:10.1242/dev.02803

Tamada Y, Yun JY, Woo SC, Amasino RM (2009) ARABIDOPSIS TRITHORAX-RELATED7 is required for methylation of lysine 4 of histone H3 and for transcriptional activation of FLOWERING LOCUS C. The Plant Cell 21 (10):3257-3269. doi:10.1105/tpc.109.070060

Tan Q, Zhang L, Grant J, Cooper P, Tegeder M (2010) Altered phloem transport of S-methylmethionine affects plant metabolism and seed number in pea plants. Plant Physiology 154:1886-1896

Tang H, Krishnakumar V, Bidwell S, Rosen B, Chan A, Zhou S, Gentzbittel L, Childs KL, Yandell M, Gundlach H, Mayer KF, Schwartz DC, Town CD (2014) An improved genome release (version Mt4.0) for the model legume Medicago truncatula. BMC Genomics 15 (1):312. doi:10.1186/1471-2164-15-312

Tank JG, Thaker VS (2011) Cyclin dependent kinases and their role in regulation of plant cell cycle. Biologia Plantarum 55 (2):201-212. doi:DOI 10.1007/s10535-011-0031-9

Taoka K-i, Ohki I, Tsuji H, Furuita K, Hayashi K, Yanase T, Yamaguchi M, Nakashima C, Purwestri YA, Tamaki S, Ogaki Y, Shimada C, Nakagawa A, Kojima C, Shimamoto K (2011) 14-3-3 proteins act as intracellular receptors for rice Hd3a florigen. Nature 476:332. doi:10.1038/nature10272

Taşan M, Musso G, Hao T, Vidal M, MacRae CA, Roth FP (2014) Selecting causal genes from genome-wide association studies via functionally coherent subnetworks. Nat Methods 12:154. doi:10.1038/nmeth.3215

Tattersall AD, Turner L, Knox MR, Ambrose MJ, Ellis THN, Hofer JMI (2005) The Mutant crispa Reveals Multiple Roles for PHANTASTICA in Pea Compound Leaf Development. The Plant Cell 17 (4):1046-1060. doi:10.1105/tpc.104.029447

Tayeh N, Aluome C, Falque M, Jacquin F, Klein A, Chauveau A, Berard A, Houtin H, Rond C, Kreplak J, Boucherot K, Martin C, Baranger A, Pilet-Nayel ML, Warkentin TD, Brunel D, Marget P, Le Paslier MC, Aubert G, Burstin J (2015a) Development of two major resources for pea genomics: the GenoPea 13.2K SNP Array and a high-density, high-resolution consensus genetic map. Plant Journal 84 (6):1257-1273. doi:10.1111/tpj.13070

Tayeh N, Aubert G, Pilet-Nayel ML, Lejeune-Henaut I, Warkentin TD, Burstin J (2015b) Genomic Tools in Pea Breeding Programs: Status and Perspectives. Frontiers in Plant Science 6. doi:ARTN 103710.3389/fpls.2015.01037

Taylor SA, Hofer JMI, Murfet IC, Sollinger JD, Singer SR, Knox MR, Ellis THN (2002) PROLIFERATING INFLORESCENCE MERISTEM, a MADS-Box Gene That Regulates Floral Meristem Identity in Pea. Plant Physiology 129 (3):1150-1159. doi:10.1104/pp.001677

Taylor SA, Murfet IC (1994) A short day mutant in pea is deficient in the floral stimulus. Flowering Newsletter (18):39-43

ten Hove CA, Lu KJ, Weijers D (2015) Building a plant: cell fate specification in the early Arabidopsis embryo. Development 142 (3):420-430. doi:10.1242/dev.111500

Teng C, Du D, Xiao L, Yu Q, Shang G, Zhao Z (2017) Mapping and Identifying a Candidate Gene (Bnmfs) for Female-Male Sterility through Whole-Genome Resequencing and RNA-Seq in Rapeseed (Brassica napus L.). Frontiers in Plant Science 8 (2086). doi:10.3389/fpls.2017.02086

Thakare D, Kumudini S, Dinkins RD (2011) The alleles at the E1 locus impact the expression pattern of two soybean FT-like genes shown to induce flowering in Arabidopsis. Planta 234 (5):933-943. doi:10.1007/s00425-011-1450-8

Theissen G (2001) Development of floral organ identity: stories from the MADS house. Current Opinion in Plant Biology 4 (1):75-85. doi:Doi 10.1016/S1369-5266(00)00139-4

Thi Tran HT, Takeshima Y, Surono A, Yagi M, Wada H, Matsuo M (2005) A G-to-A transition at the fifth position of intron-32 of the dystrophin gene inactivates a splice-donor site both in vivo and in vitro. Molecular Genetics and Metabolism 85 (3):213-219. doi:https://doi.org/10.1016/j.ymgme.2005.03.006

Thole JM, Strader LC (2015) Next-generation sequencing as a tool to quickly identify causative EMS-generated mutations. Plant Signaling & Behavior 10 (5):1-4. doi:10.1080/15592324.2014.1000167

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research 25 (24):4876-4882

Thorstensen T, Grini PE, Aalen RB (2011) SET domain proteins in plant development. Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms 1809 (8):407-420. doi:https://doi.org/10.1016/j.bbagrm.2011.05.008

Tilbrook K, Arongaus AB, Binkert M, Heijde M, Yin R, Ulm R (2013) The UVR8 UV-B Photoreceptor: Perception, Signaling and Response. The Arabidopsis Book / American Society of Plant Biologists 11:e0164. doi:10.1199/tab.0164

Tivendale ND, Davidson SE, Davies NW, Smith JA, Dalmais M, Bendahmane AI, Quittenden LJ, Sutton L, Bala RK, Le Signor C, Thompson R, Horne J, Reid JB, Ross JJ (2012) Biosynthesis of the Halogenated Auxin, 4-Chloroindole-3-Acetic Acid. Plant Physiology 159 (3):1055-1063. doi:10.1104/pp.112.198457

Tiwari SB, Shen Y, Chang H, Hou Y, Harris A, Ma SF, McPartland M, Hymus GJ, Adam L, Marion C, Belachew A, Repetti PP, Reuber TL, Ratcliffe OJ (2010) The flowering time regulator CONSTANS is recruited to the FLOWERING LOCUS T promoter via a unique cis-element. New Phytologist 187 (1):57-66. doi:10.1111/j.1469-8137.2010.03251.x

Todorov DT, Karanov EN, Smith AR, Hall MA (2003) Chlorophyllase activity and chlorophyll content in wild type and eti 5 mutant of Arabidopsis thaliana subjected to low and high temperatures. Biologia Plantarum 46 (4):633-636. doi:Doi 10.1023/A:1024896418839

Tooke F, Ordidge M, Chiurugwi T, Battey N (2005) Mechanisms and function of flower and inflorescence reversion. Journal of Experimental Botany 56 (420):2587-2599. doi:10.1093/jxb/eri254

Tooke F, Pouteau S, Battey N (1998) Non-reversion of Impatiens in the absence of meristem commitment. Journal of Experimental Botany 49 (327):1681-1688. doi:DOI 10.1093/jexbot/49.327.1681

Trevaskis B, Hemming MN, Dennis ES, Peacock WJ (2007) The molecular basis of vernalization-induced flowering in cereals. Trends in Plant Science 12 (8):352-357. doi:10.1016/j.tplants.2007.06.010

Triques K, Sturbois B, Gallais S, Dalmais M, Chauvin S, Clepet C, Aubourg S, Rameau C, Caboche M, Bendahmane A (2007) Characterization of Arabidopsis thaliana mismatch specific endonucleases: application to mutation discovery by TILLING in pea. The Plant Journal 51 (6):1116-1125. doi:doi:10.1111/j.1365-313X.2007.03201.x

Tsai K-L, Sato S, Tomomori-Sato C, Conaway RC, Conaway JW, Asturias FJ (2013) A conserved Mediator–CDK8 kinase module association regulates Mediator–RNA polymerase II interaction. Nature Structural & Molecular Biology 20 (5):611-619. doi:10.1038/nsmb.2549

Tschiersch B, Hofmann A, Krauss V, Dorn R, Korge G, Reuter G (1994) The Protein Encoded by the Drosophila Position-Effect Variegation Suppressor Gene Su(Var)3-9 Combines Domains of Antagonistic Regulators of Homeotic Gene Complexes. EMBO Journal 13 (16):3822-3831

Tsuji H, Taoka K, Shimamoto K (2013) Florigen in rice: complex gene network for florigen transcription, florigen activation complex, and multiple functions. Current Opinion in Plant Biology 16 (2):228-235. doi:10.1016/j.pbi.2013.01.005

Turck F, Roudier F, Farrona S, Martin-Magniette ML, Guillaume E, Buisine N, Gagnot S, Martienssen RA, Coupland G, Colot V (2007) Arabidopsis TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. PLOS Genetics 3 (6):855-866. doi:ARTN e8610.1371/journal.pgen.0030086

Turner A, Beales J, Faure S, Dunford RP, Laurie DA (2005) The pseudo-response regulator Ppd-H1 provides adaptation to photoperiod in barley. Science 310 (5750):1031-1034. doi:10.1126/science.1117619

van Daal A, White EM, Elgin SCR, Gorovsky MA (1990) Conservation of intron position indicates separation of major and variant H2As is an early event in the evolution of eukaryotes. Journal of Molecular Evolution 30 (5):449-455. doi:10.1007/BF02101116

Vasyukova NI, Chalenko GI, Gerasimova NG, Ozeretskovskaya OL (2011) Wound repair in plant tissues (Review). Applied Biochemistry and Microbiology 47 (3):229-233. doi:10.1134/s0003683811030161

Ventura Y, Myrzabayeva M, Alikulov Z, Omarov R, Khozin-Goldberg I, Sagi M (2014) Effects of salinity on flowering, morphology, biomass accumulation and leaf metabolites in an edible halophyte. AoB PLANTS 6:plu053-plu053. doi:10.1093/aobpla/plu053

Vilela C, McCarthy JEG (2003) Regulation of fungal gene expression via short open reading frames in the mRNA 5′ untranslated region. Molecular Microbiology 49 (4):859-867. doi:10.1046/j.1365-2958.2003.03622.x

von Bohlen AE, Bohm J, Pop R, Johnson DS, Tolmie J, Stucker R, Morris-Rosendahl D, Scherer G (2017) A mutation creating an upstream initiation codon in the SOX9 5' UTR causes acampomelic campomelic dysplasia. Molecular Genetics & Genomic Medicine 5 (3):261-268. doi:10.1002/mgg3.282

Wang DG, Fan JB, Siao CJ, Berno A, Young P, Sapolsky R, Ghandour G, Perkins N, Winchester E, Spencer J, Kruglyak L, Stein L, Hsie L, Topaloglou T, Hubbell E, Robinson E, Mittmann M, Morris MS, Shen NP, Kilburn D, Rioux J, Nusbaum C, Rozen S, Hudson TJ, Lipshutz R, Chee M, Lander ES (1998) Large-scale identification, mapping, and genotyping of single-nucleotide polymorphisms in the human genome. Science 280 (5366):1077-1082. doi:DOI 10.1126/science.280.5366.1077

Wang GF, Kong HZ, Sun YJ, Zhang XH, Zhang W, Altman N, dePamphilis CW, Ma H (2004) Genome-wide analysis of the cyclin family in Arabidopsis and comparative phylogenetic analysis of plant cyclin-like proteins. Plant Physiology 135 (2):1084-1099. doi:10.1104/pp.104.04036

Wang H, Pan J, Li Y, Lou D, Hu Y, Yu D (2016) The DELLA-CONSTANS Transcription Factor Cascade Integrates Gibberellic Acid and Photoperiod Signaling to Regulate Flowering. Plant Physiology 172 (1):479-488. doi:10.1104/pp.16.00891

Wang S, Cao M, Ma X, Chen W, Zhao J, Sun C, Tan L, Liu F (2017) Integrated RNA Sequencing and QTL Mapping to Identify Candidate Genes from Oryza rufipogon Associated with Salt Tolerance at the Seedling Stage. Frontiers in Plant Science 8 (1427). doi:10.3389/fpls.2017.01427

Wang W, Chen X (2004) HUA ENHANCER3 reveals a role for a cyclin-dependent protein kinase in the specification of floral organ identity in Arabidopsis. Development (Cambridge, England) 131 (13):3147-3156. doi:10.1242/dev.01187

Wang X, Wang J, Ding Z, Ji J, Sun Q, Cai G (2013) Structural flexibility and functional interaction of mediator Cdk8 module. Protein & Cell 4 (12):911-920. doi:10.1007/s13238-013-3069-y

Wang YAN, Liu J, Huang BO, Xu Y-M, Li J, Huang L-F, Lin JIN, Zhang J, Min Q-H, Yang W-M, Wang X-Z (2015) Mechanism of alternative splicing and its regulation. Biomedical Reports 3 (2):152-158. doi:10.3892/br.2014.407

Wang Z, Luo Y, Li X, Wang L, Xu S, Yang J, Weng L, Sato S, Tabata S, Ambrose M, Rameau C, Feng X, Hu X, Luo D (2008) Genetic control of floral zygomorphy in pea (Pisum sativum L.). Proceedings of the National Academy of Sciences 105 (30):10414-10419. doi:10.1073/pnas.0803291105

Wargent JJ, Gegas VC, Jenkins GI, Doonan JH, Paul ND (2009a) UVR8 in Arabidopsis thaliana regulates multiple aspects of cellular differentiation during leaf development in response to ultraviolet B radiation. New Phytologist 183 (2):315-326. doi:10.1111/j.1469-8137.2009.02855.x

Wargent JJ, Moore JP, Roland Ennos A, Paul ND (2009b) Ultraviolet Radiation as a Limiting Factor in Leaf Expansion and Development. Photochemistry and Photobiology 85 (1):279-286. doi:10.1111/j.1751-1097.2008.00433.x

Watanabe S, Harada K, Abe J (2012) Genetic and molecular bases of photoperiod responses of flowering in soybean. Breeding Science 61 (5):531-543. doi:10.1270/jsbbs.61.531

Watanabe S, Hideshima R, Xia Z, Tsubokura Y, Sato S, Nakamoto Y, Yamanaka N, Takahashi R, Ishimoto M, Anai T, Tabata S, Harada K (2009) Map-based cloning of the gene associated with the soybean maturity locus E3. Genetics 182 (4):1251-1262. doi:10.1534/genetics.108.098772

Wathugala DL, Hemsley PA, Moffat CS, Cremelie P, Knight MR, Knight H (2012) The Mediator subunit SFR6/MED16 controls defence gene expression mediated by salicylic acid and jasmonate responsive pathways. New Phytologist 195 (1):217-230. doi:10.1111/j.1469-8137.2012.04138.x

Weller J (2007) Update on the genetics of flowering. Pisum Genetics 39:1-8

Weller J, Murfet I, Reid J (1992) Internode length in Pisum. Two further lv mutants. Pisum Genetics 24:86-89

Weller JL, Batge SL, Smith JJ, Kerckhoffs LH, Sineshchekov VA, Murfet IC, Reid JB (2004) A dominant mutation in the pea PHYA gene confers enhanced responses to light and impairs the light-dependent degradation of phytochrome A. Plant Physiology 135 (4):2186-2195. doi:10.1104/pp.103.036103

Weller JL, Beauchamp N, Kerckhoffs LHJ, Platten JD, Reid JB (2001) Interaction of phytochromes A and B in the control of de-etiolation and flowering in pea. The Plant Journal 26 (3):283-294. doi:10.1046/j.1365-313X.2001.01027.x

Weller JL, Foo EM, Hecht V, Ridge S, Vander Schoor JK, Reid JB (2015) Ethylene Signaling Influences Light-Regulated Development in Pea. Plant Physiology 169 (1):115-124. doi:10.1104/pp.15.00164

Weller JL, Hecht V, Schoor JKV, Davidson SE, Ross JJ (2009) Light Regulation of Gibberellin Biosynthesis in Pea Is Mediated through the COP1/HY5 Pathway. The Plant Cell 21 (3):800-813. doi:10.1105/tpc.108.063628

Weller JL, Hecht VFG, Sussmilch FC (2013) Isolation and Forward Genetic Analysis of Developmental Genes in Pea. In: Rose RJ (ed) Legume Genomics: Methods and Protocols. Humana Press, Totowa, NJ, pp 147-161. doi:10.1007/978-1-62703-613-9_11

Weller JL, Liew LC, Hecht VF, Rajandran V, Laurie RE, Ridge S, Wenden B, Vander Schoor JK, Jaminon O, Blassiau C, Dalmais M, Rameau C, Bendahmane A, Macknight RC, Lejeune-Henaut I (2012a) A conserved molecular basis for photoperiod adaptation in two temperate legumes. Proceedings of the National Academy of Sciences of the United States of America 109 (51):21158-21163. doi:10.1073/pnas.1207943110

Weller JL, Liew LC, Hecht VFG, Rajandran V, Laurie RE, Ridge S, Wenden B, Vander Schoor JK, Jaminon O, Blassiau C, Dalmais M, Rameau C, Bendahmane A, Macknight RC, Lejeune-Hénaut I (2012b) A conserved molecular basis for photoperiod adaptation in two temperate legumes. Proceedings of the National Academy of Sciences 109 (51):21158-21163. doi:10.1073/pnas.1207943110

Weller JL, Murfet IC, Reid JB (1997a) Pea mutants with reduced sensitivity to far-red light define an important role for phytochrome A in day-length detection (vol 114, pg 1225, 1997). Plant Physiology 115 (3):1291-1291

Weller JL, Nagatani A, Kendrick RE, Murfet IC, Reid JB (1995) New lv Mutants of Pea Are Deficient in Phytochrome B. Plant Physiology 108 (2):525-532

Weller JL, Ortega R (2015) Genetic control of flowering time in legumes. Frontiers in Plant Science 6 : 207. Weller, J. L., & Ortega, R. (2015). Genetic control of flowering time in legumes. Frontiers in plant science, 6, 207. doi:10.3389/fpls.2015.00207

Weller JL, Reid JB (1993) Photoperiodism and Photocontrol of Stem Elongation in 2 Photomorphogenic Mutants of Pisum-Sativum L. Planta 189 (1):15-23

Weller JL, Reid JB, Taylor SA, Murfet IC (1997b) The genetic control of flowering in pea. Trends in Plant Science 2 (11):412-418

Wessagowit V, Nalla VK, Rogan PK, McGrath JA (2005) Normal and abnormal mechanisms of gene splicing and relevance to inherited skin diseases. Journal of Dermatological Science 40 (2):73-84. doi:10.1016/j.jdermsci.2005.05.006

Westermarck J, Ivaska J, Corthals GL (2013) Identification of Protein Interactions Involved in Cellular Signaling. Molecular & Cellular Proteomics : MCP 12 (7):1752-1763. doi:10.1074/mcp.R113.027771

Weston DE, Elliott RC, Lester DR, Rameau C, Reid JB, Murfet IC, Ross JJ (2008) The pea DELLA proteins LA and CRY are important regulators of gibberellin synthesis and root growth. Plant Physiology 147 (1):199-205. doi:10.1104/pp.108.115808

Weston DE, Reid JB, Ross JJ (2009) Auxin regulation of gibberellin biosynthesis in the roots of pea (Pisum sativum). Functional Plant Biology 36 (4):362-369. doi:10.1071/Fp08301

Wethmar K, Barbosa-Silva A, Andrade-Navarro MA, Leutz A (2014) uORFdb-a comprehensive literature database on eukaryotic uORF biology. Nucleic Acids Research 42 (D1):D60-D67. doi:10.1093/nar/gkt952

White OE (1917) Studies of Inheritance in Pisum. II. The Present State of Knowledge of Heredity and Variation in Peas. Proceedings of the American Philosophical Society 56 (7):487-588

Whitelam GC, Devlin PF (1997) Roles of different phytochromes in Arabidopsis photomorphogenesis. Plant, Cell & Environment 20 (6):752-758. doi:doi:10.1046/j.1365-3040.1997.d01-100.x

Whittaker C, Dean C (2017a) The FLC Locus: A Platform for Discoveries in Epigenetics and Adaptation. Annual Review of Cell and Developmental Biology 33 (1):555-575. doi:10.1146/annurev-cellbio-100616-060546

Wigge PA (2013) Ambient temperature signalling in plants. Current Opinion in Plant Biology 16 (5):661-666. doi:10.1016/j.pbi.2013.08.004

Wigge PA, Kim MC, Jaeger KE, Busch W, Schmid M, Lohmann JU, Weigel D (2005) Integration of spatial and temporal information during floral induction in Arabidopsis. Science 309 (5737):1056-1059. doi:DOI 10.1126/science.1114358

Wilson ZA, Yang C (2004) Plant gametogenesis: conservation and contrasts in development. Reproduction 128 (5):483-492. doi:10.1530/rep.1.00306

Wiltshire RJE, Murfet IC, Reid JB (1994) The genetic control of heterochrony: Evidence from developmental mutants of Pisum sativum L. Journal of Evolutionary Biology 7 (4):447-465. doi:doi:10.1046/j.1420-9101.1994.7040447.x

Wong ACS, Hecht VFG, Picard K, Diwadkar P, Laurie RE, Wen J, Mysore K, Macknight RC, Weller JL (2014) Isolation and functional analysis of CONSTANS-LIKE genes suggests that a central role for CONSTANS in flowering time control is not evolutionarily conserved in Medicago truncatula. Frontiers in Plant Science 5. doi:10.3389/fpls.2014.00486

Wood CC, Robertson M, Tanner G, Peacock WJ, Dennis ES, Helliwell CA (2006) The Arabidopsis thaliana vernalization response requires a polycomb-like protein complex that also includes VERNALIZATION INSENSITIVE 3. Proceedings of the National Academy of Sciences of the United States of America 103 (39):14631-14636. doi:10.1073/pnas.0606385103

Worland AJ, Börner A, Korzun V, Li WM, Petrovíc S, Sayers EJ (1998) The influence of photoperiod genes on the adaptability of European winter wheats. Euphytica 100 (1-3):385-394

Wu F, Price BW, Haider W, Seufferheld G, Nelson R, Hanzawa Y (2014) Functional and evolutionary characterization of the CONSTANS gene family in short-day photoperiodic flowering in soybean. PLOS One 9 (1). doi:10.1371/journal.pone.0085754

Xia Z, Watanabe S, Yamada T, Tsubokura Y, Nakashima H, Zhai H, Anai T, Sato S, Yamazaki T, Lü S, Wu H, Tabata S, Harada K (2012) Positional cloning and characterization reveal the molecular basis for soybean maturity locus E1 that regulates photoperiodic flowering. Proceedings of the National Academy of Sciences of the United States of America 109 (32):E2155-E2164. doi:10.1073/pnas.1117982109

Xiao ZL, Hu Y, Zhang XL, Xue YQ, Fang ZY, Yang LM, Zhang YY, Liu YM, Li ZS, Liu X, Liu ZZ, Lv HH, Zhuang M (2017) Fine Mapping and Transcriptome Analysis Reveal Candidate Genes Associated with Hybrid Lethality in Cabbage (Brassica Oleracea). Genes-Basel 8 (6). doi:ARTN 14710.3390/genes8060147

Xie Q, Wang P, Liu X, Yuan L, Wang L, Zhang C, Li Y, Xing H, Zhi L, Yue Z, Zhao C, McClung CR, Xu X (2014) LNK1 and LNK2 are transcriptional coactivators in the Arabidopsis circadian oscillator. The Plant Cell 26 (7):2843-2857. doi:10.1105/tpc.114.126573

Xie X-J, Hsu F-N, Gao X, Xu W, Ni J-Q, Xing Y, Huang L, Hsiao H-C, Zheng H, Wang C, Zheng Y, Xiaoli AM, Yang F, Bondos SE, Ji J-Y (2015) CDK8-Cyclin C Mediates Nutritional Regulation of Developmental Transitions through the Ecdysone Receptor in Drosophila. PLOS Biology 13 (7):e1002207. doi:10.1371/journal.pbio.1002207

Xing DH, Zhao HW, Xu RQ, Li QSQ (2008) Arabidopsis PCFS4, a homologue of yeast polyadenylation factor Pcf11p, regulates FCA alternative processing and promotes flowering time. Plant Journal 54 (5):899-910. doi:10.1111/j.1365-313X.2008.03455.x

Xu L, Zhao Z, Dong A, Soubigou-Taconnat L, Renou JP, Steinmetz A, Shen WH (2008) Di- and tri- but not monomethylation on histone H3 lysine 36 marks active transcription of genes involved in flowering time regulation and other processes in Arabidopsis thaliana. Molecular and Cellular Biology 28 (4):1348-1360. doi:10.1128/MCB.01607-07

Xue WY, Xing YZ, Weng XY, Zhao Y, Tang WJ, Wang L, Zhou HJ, Yu SB, Xu CG, Li XH, Zhang QF (2008) Natural variation in Ghd7 is an important regulator of heading date and yield potential in rice. Nature Genetics 40 (6):761-767. doi:DOI 10.1038/ng.143

Yadav SK (2010) Cold stress tolerance mechanisms in plants. A review. Agronomy for Sustainable Development 30 (3):515-527. doi:10.1051/agro/2009050

Yamada K, Fukao Y, Hayashi M, Fukazawa M, Suzuki I, Nishimura M (2007) Cytosolic HSP90 regulates the heat shock response that is responsible for heat acclimation in Arabidopsis thaliana. The Journal of Biological Chemistry 282 (52):37794-37804. doi:10.1074/jbc.M707168200

Yamaguchi A, Wu MF, Yang L, Wu G, Poethig RS, Wagner D (2009) The MicroRNA-Regulated SBP-Box Transcription Factor SPL3 Is a Direct Upstream Activator of LEAFY, FRUITFULL, and APETALA1. Developmental Cell 17 (2):268-278. doi:DOI 10.1016/j.devcel.2009.06.007

Yamaguchi T, Hamamoto S, Uozumi N (2013) Sodium transport system in plant cells. Frontiers in Plant Science 4:410. doi:10.3389/fpls.2013.00410

Yan L, Fu D, Li C, Blechl A, Tranquilli G, Bonafede M, Sanchez A, Valarik M, Yasuda S, Dubcovsky J (2006) The wheat and barley vernalization gene VRN3 is an orthologue of FT. Proceedings of the National Academy of Sciences of the United States of America 103 (51):19581-19586. doi:10.1073/pnas.0607142103

Yan L, Loukoianov A, Blechl A, Tranquilli G, Ramakrishna W, SanMiguel P, Bennetzen JL, Echenique V, Dubcovsky J (2004) The wheat VRN2 gene is a flowering repressor down-regulated by vernalization. Science 303 (5664):1640-1644. doi:10.1126/science.1094305

Yan L, Loukoianov A, Tranquilli G, Helguera M, Fahima T, Dubcovsky J (2003) Positional cloning of the wheat vernalization gene VRN1. Proceedings of the National Academy of Sciences of the United States of America 100 (10):6263-6268. doi:10.1073/pnas.0937399100

Yang L, Liu T, Li B, Sui Y, Chen J, Shi J, Wing RA, Chen M (2012) Comparative Sequence Analysis of the Ghd7 Orthologous Regions Revealed Movement of Ghd7 in the Grass Genomes. PLOS One 7 (11):e50236. doi:10.1371/journal.pone.0050236

Yang S, Murphy RL, Morishige DT, Klein PE, Rooney WL, Mullet JE (2014a) Sorghum phytochrome B inhibits flowering in long days by activating expression of SbPRR37 and SbGHD7, repressors of SbEHD1, SbCN8 and SbCN12. PLOS One 9 (8). doi:10.1371/journal.pone.0105352

Yang S, Weers BD, Morishige DT, Mullet JE (2014b) CONSTANS is a photoperiod regulated activator of flowering in sorghum. BMC Plant Biology 14 (1). doi:10.1186/1471-2229-14-148

Yang Y, Li L, Qu L-J (2016) Plant Mediator complex and its critical functions in transcription regulation. Journal of integrative plant biology 58 (2):106-118. doi:10.1111/jipb.12377

Yang YQ, Yao YN, He H (2008) Influence of ambient and enhanced ultraviolet-B radiation on the plant growth and physiological properties in two contrasting populations of Hippophae rhamnoides. Journal of Plant Research 121 (4):377-385. doi:10.1007/s10265-008-0163-y

Yoshida S, Barbier de Reuille P, Lane B, Bassel GW, Prusinkiewicz P, Smith RS, Weijers D (2014) Genetic control of plant development by overriding a geometric division rule. Developmental Cell 29 (1):75-87. doi:10.1016/j.devcel.2014.02.002

Yoshida T, Ohama N, Nakajima J, Kidokoro S, Mizoi J, Nakashima K, Maruyama K, Kim JM, Seki M, Todaka D, Osakabe Y, Sakuma Y, Schoffl F, Shinozaki K, Yamaguchi-Shinozaki K (2011) Arabidopsis HsfA1 transcription factors function as the main positive regulators in heat shock-responsive gene expression. Molecular genetics and genomics : MGG 286 (5-6):321-332. doi:10.1007/s00438-011-0647-7

Young ND, Udvardi M (2009) Translating Medicago truncatula genomics to crop legumes. Current Opinion in Plant Biology 12 (2):193-201. doi:10.1016/j.pbi.2008.11.005

Yu CW, Liu X, Luo M, Chen C, Lin X, Tian G, Lu Q, Cui Y, Wu K (2011a) HISTONE DEACETYLASE6 interacts with FLOWERING LOCUS D and regulates flowering in Arabidopsis. Plant Physiology 156 (1):173-184. doi:10.1104/pp.111.174417

Yu S, Galvao VC, Zhang YC, Horrer D, Zhang TQ, Hao YH, Feng YQ, Wang S, Schmid M, Wang JW (2012) Gibberellin regulates the Arabidopsis floral transition through miR156-targeted SQUAMOSA promoter binding-like transcription factors. The Plant Cell 24 (8):3320-3332. doi:10.1105/tpc.112.101014

Yu X, Li L, Zola J, Aluru M, Ye H, Foudree A, Guo H, Anderson S, Aluru S, Liu P (2011b) A brassinosteroid transcriptional network revealed by genome‐wide identification of BESI target genes in Arabidopsis thaliana. The Plant Journal 65 (4):634-646

Yuan JS, Galbraith DW, Dai SY, Griffin P, Stewart CN (2008a) Plant systems biology comes of age. Trends in Plant Science 13 (4):165-171. doi:10.1016/j.tplants.2008.02.003

Yuan JS, Tiller KH, Al-Ahmad H, Stewart NR, Stewart CN (2008b) Plants to power: bioenergy to fuel the future. Trends in Plant Science 13 (8):421-429. doi:10.1016/j.tplants.2008.06.001

Yun J-Y, Tamada Y, Kang YE, Amasino RM (2012) ARABIDOPSIS TRITHORAX-RELATED3/SET DOMAIN GROUP2 is Required for the Winter-Annual Habit of Arabidopsis thaliana. Plant and Cell Physiology 53 (5):834-846. doi:10.1093/pcp/pcs021

Zakhrabekova S, Gough SP, Braumann I, MuÏller AH, Lundqvist J, Ahmann K, Dockter C, Matyszczak I, Kurowska M, Druka A, Waugh R, Granerd A, Stein N, Steuernagel B, Lundqvist U, Hansson M (2012) Induced mutations in circadian clock regulator Mat-a facilitated short-season adaptation and range extension in cultivated barley. Proceedings of the National Academy of Sciences of the United States of America 109 (11):4326-4331. doi:10.1073/pnas.1113009109

Zandalinas SI, Balfagón D, Arbona V, Gómez-Cadenas A, Inupakutika MA, Mittler R (2016) ABA is required for the accumulation of APX1 and MBF1c during a combination of water deficit and heat stress. Journal of Experimental Botany 67 (18):5381-5390. doi:10.1093/jxb/erw299

Zhai H, Lü S, Liang S, Wu H, Zhang X, Liu B, Kong F, Yuan X, Li J, Xia Z (2014) GmFT4, a homolog of FLOWERING LOCUS T, is positively regulated by E1 and functions as a flowering repressor in soybean. PLOS One 9 (2). doi:10.1371/journal.pone.0089030

Zhang LZ, Garneau MG, Majumdar R, Grant J, Tegeder M (2015) Improvement of pea biomass and seed productivity by simultaneous increase of phloem and embryo loading with amino acids. Plant Journal 81 (1):134-146. doi:10.1111/tpj.12716

Zhang LZ, Tan QM, Lee R, Trethewy A, Lee YH, Tegeder M (2010) Altered Xylem-Phloem Transfer of Amino Acids Affects Metabolism and Leads to Increased Seed Yield and Oil Content in Arabidopsis. The Plant Cell 22 (11):3603-3620. doi:10.1105/tpc.110.073833

Zhang XY, Clarenz O, Cokus S, Bernatavichute YV, Pellegrini M, Goodrich J, Jacobsen SE (2007) Whole-genome analysis of histone H3 lysine 27 trimethylation in Arabidopsis. PLOS Biology 5 (5):1026-1035. doi:ARTN e12910.1371/journal.pbio.0050129

Zhang XY, Yazaki J, Sundaresan A, Cokus S, Chan SWL, Chen HM, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE, Ecker JR (2006) Genome-wide high-resolution mapping and functional analysis of DNA methylation in Arabidopsis. Cell 126 (6):1189-1201. doi:10.1016/j.cell.2006.08.003

Zhao XY, Liu MS, Li JR, Guan CM, Zhang XS (2005) The wheat TaGI1, involved in photoperiodic flowering, encodes an Arabidopsis GI ortholog. Plant Molecular Biology 58 (1):53-64. doi:10.1007/s11103-005-4162-2

Zheng B, Chen X (2011) Dynamics of histone H3 lysine 27 trimethylation in plant development. Current Opinion in Plant Biology 14 (2):123-129. doi:10.1016/j.pbi.2011.01.001

Zheng Z, Guan H, Leal F, Grey PH, Oppenheimer DG (2013) Mediator Subunit18 Controls Flowering Time and Floral Organ Identity in Arabidopsis. PLOS One 8 (1):e53924. doi:10.1371/journal.pone.0053924

Zhu J, Mayeda A, Krainer AR (2001) Exon Identity Established through Differential Antagonism between Exonic Splicing Silencer-Bound hnRNP A1 and Enhancer-Bound SR Proteins. Molecular Cell 8 (6):1351-1361. doi:https://doi.org/10.1016/S1097-2765(01)00409-9

Zhu Y, Schluttenhoffer CM, Wang P, Fu F, Thimmapuram J, Zhu J-K, Lee SY, Yun D-J, Mengiste T (2014) CYCLIN-DEPENDENT KINASE8 Differentially Regulates Plant Immunity to Fungal Pathogens through Kinase-Dependent and -Independent Functions in Arabidopsis. The Plant Cell 26 (10):4149-4170. doi:10.1105/tpc.114.128611

Zhu YL, Song QJ, Hyten DL, Van Tassell CP, Matukumalli LK, Grimm DR, Hyatt SM, Fickus EW, Young ND, Cregan PB (2003) Single-nucleotide polymorphisms in soybean. Genetics 163 (3):1123-1134

Zografou T, Turck F (2013) Epigenetic Control of Flowering Time. In: Grafi G, Ohad N (eds) Epigenetic Memory and Control in Plants, vol 18. Signaling and Communication in Plants. Springer Berlin Heidelberg, pp 77-105. doi:10.1007/978-3-642-35227-0_5

Zuo Z, Liu H, Liu B, Liu X, Lin C (2011) Blue light-dependent interaction of CRY2 with SPA1 regulates COP1 activity and floral initiation in Arabidopsis. Current Biology : CB 21 (10):841-847. doi:10.1016/j.cub.2011.03.048

Appendix 1 – Details of primers

Table A1.1. Details on all the PCRs that were used in the attempt of finding intron based polymorphism between *late3-1* and cv. Terese

Locus name	*M. truncatula* accession	*P. sativum* accession	Primer name	Primer sequence	Tm (° C)	Intron number	cDNA (bp)	PCR product	Sequ-enced	Polym-orphic	Map-ped
ATNAP13	Medtr3g096300	PsCam034645	PsATNAP13_1F	TGAGTTTTACCGCTCAGC	55	2	421	No			
			PsATNAP13_1R	TGTAAAGCTTGGGACACC	54.4						
ATX6	Medtr3g095840	PsCam038017	PsATX6-1F	GTTCACGCAAAGGAGAATACTTG	58.3	1	151	No			
			PsATX6-1R	GGGCAGTAACCTTCTCCTTTTC	58.9						
			PsATX6-2F	ACCCTGCTACCAGTCTTGTTG	59.9	1	149	500 bp	Yes	Yes	Yes
			PsATX6-2R	AAAGACCCGCGATTCATCATTAAC	60						
BTB1	Medtr3g096160	PsCam000088	PsBTB1_1F	TTCCCGAAGTTGATAATTTGGG	56.7	1	193	800 bp	Yes	Yes	Yes
			PsBTB1_1R	TCTTCTGTCCAAGGAACAGC	57.4						
CDDP1	Medtr3g095470	PsCam037008	PsCDDP1_1F	GATGGCGAAATTTTGTGTCC	55.9	3	294	No			
			PsCDDP1_1R	TGGGATGGCTATTTGAACG	55.3						
	Medtr3g095470	PsCam037008	PsCDDP1_2F	TCTGGTCGACGATTTGGTGG	60	1	103	300 bp	Yes	No	
			PsCDDP1_2R	TCTGTTCAGGATGGCCATCG	59.8						
COX15	Medtr3g095790	PsCam043034	PsCOX15_1F	GAAATTTACGGGTGAGTTCC	54.3	2	367	No			
			PsCOX15_1R	AGCTGCAAGACGATAAGG	54.4						
CR88	Medtr3g096860	PsCam021158	PsCR88_1F	CGGCAAGCTTCTCATGG	55.6	3	257	No			
			PsCR88_1R	CCAGATGAATGTGTAACTTGG	54.6						
	Medtr3g096860	PsCam021158	PsCR88_2F	TACCGGCAAGCTTCTCATGG	60.1	1	108	600 bp	Yes	Yes	No
			PsCR88_2R	TCAGAACTGGTGAAGACTCTCC	59.1						
DEK1	Medtr3g098990	PsCam016767	PsDEK1_1F	GACCAGAGAACGGAAGG	53.6	4	313	No			
			PsDEK1_1R	CCTTTCTTTTGCCTTTATCC	52.4						
DYAD1	Medtr3g097570	PsCam049636	PsDYAD1_1F	GGTGAAAGAGGAGTCATGC	55.6	1	340	450 bp	Yes	No	
			PsDYAD1_1R	TCCAGGTGCCACTTTACC	56.4						
G20	Medtr3g096500	PsCam046205	PsG20_1F	TGGTGGAATCTCTGGATACTCC	58.7	1	244	2200 bp	Yes	Yes	No
			PsG20_1R	CCCCAAACTAATGGCCAATAGC	59.6						
LCV2	Medtr3g095210	PsCam021197	PsLCV2_1F	TCAATTCCTCTGAGTCAAGC	55.1	3	328	No			
			PsLCV2_1R	CCAATCCAGAAGACAGTGG	55.2						

Table A1.1 (Continued)

Locus name	*M. truncatula* accession	*P. sativum* accession	Primer name	Primer sequence	Tm (° C)	Intron number	cDNA (bp)	PCR product	Sequ-enced	Polym-orphic	Map-ped
	Medtr3g095210	PsCam021197	PsLCV2_2F	TCTATTCCCAGTTGCAGTTACC	57.8	1	173	700 bp	Yes	Yes	Yes
			PsLCV2_2R	CCAATCCAGAAGACAGTGGC	58.5						
NST1	Medtr3g095130	PsCam023354	PsNST1_1F	AGCATCATGGTTGTTTAACG	54.6	2	383	No			
			PsNST1_1R	CAACACCAGCCAAAACC	53.8						
OSBP1	Medtr3g097270	PsCam059306	PsOSBP1_1F	GGTCGTTCGATTCAGCTCG	58.7	1	154	250 bp	Yes	No	
			PsOSBP1_1R	TCCCCTGTATGTTCATTGTACC	57.5						
PPT1	Medtr3g098060	PsCam035950	PsPPT1_1F	CACCCTCGTGTTGTGTATTGC	59.8	1	147	250 bp	Yes	No	
			PsPPT1_1R	TTACGCGCCCAAATGTCC	58.4						
SCPL45	Medtr3g498755	PsCam034080	PsSCPL45_1F	GACACCGATTGGAGTTGG	55.4	3	444	No			
			PsSCPL45_1R	TCTTTGCTCACCAATCCC	54.4						
SPS1	Medtr3g098260	PsCam045416	PsSPS1_1F	AGGATGCAGCAAGAGGAACG	60.4	1	180	300 bp	Yes	Yes	Yes
			PsSPS1_1R	AGCAATGCTCTCCTCGTGG	59.8						
TPR1	Medtr3g095070	PsCam038860	PsTPR1_1F	ACCGAATGAAGTTGAGTGG	55.1	1	201	600 bp	Yes	Yes	Yes
			PsTPR1_1R	AACTCTAGAGCTTTGTTCAACG	56.7						
WRKY13	Medtr3g095080	PsCam039266	PsWRKY13_1F	CTGAAGCTGCAACATTCG	54.6	1	227	No			
			PsWRKY13_1R	CTTTATGGCTGCTGAAACC	54.4						

Table A1.2. Details on all the PCRs that were used in the attempt of finding intron based polymorphism between *late4-1* and cv. Terese.

Locus name	*M. truncatula* accession	*P. sativum* accession	Primer name	Primer sequence	Tm (° C)	Intron number	cDNA (bp)	PCR product	Sequ-enced	Polym-orphic	Map-ped
ABA4	Medtr7g007280	PsCam001816	PsABA4_1F	AATGGTGTCGAGCTATGC	54.8	2	339	No			
			PsABA4_1R	TTCAGGTGTCCAAGAAAGG	54.6						
			PsABA4_2F	ACAACATCGTTTCGCAATCG	57.8	1	106	300 bp	Yes	Yes	Yes
			PsABA4_2R	GGAGCACTGCGATTGTTCC	59.2						
ABHP1	Medtr7g007310	PsCam038878	PsABHP1_1F	AGTTACCATGCTGCTTGG	54.9	2	189	400 bp	Yes	yes	yes
			PsABHP1_1R	TGGCGATTGAACGTTTCG	56.5						

Table A1.2. (continued)

Locus name	*M. truncatula* accession	*P. sativum* accession	Primer name	Primer sequence	Tm (° C)	Intron number	cDNA (bp)	PCR product	Sequ-enced	Polym-orphic	Map-ped
ADP1	Medtr7g020860	PsCam035619	PsADP1-1F	TGGCGTTCTTGTTTGCATTG	58.2	1	170	300 bp	Yes	No	
			PsADP1-1R	GCAAAAGTTCCTCCCAGACTG	59.1						
ATPB1	Medtr7g055720	PsCam033465	PsATPB1-1F	GACATACTCTTGAGCATGCTGATAG	59.4	1	141	300 bp	Yes	Yes	Yes
			PsATPB1-1R	ACTTATTTTGAGCCAACCCAATG	57.8						
ATR3	Medtr7g406970	PsCam057504	PsATR3_1F	TTGGATGCTGCTGAACG	55.1	4	450	No			
			PsATR3_1R	AACATCTGGGCCATTTGG	55.1						
BHLH1	Medtr7g053410	PsCam042714	PsBHLH1-1F	TAATGGAACTATGCCCCACAACTC	60.6	1	160	No			
			PsBHLH1-1R	TCCCATTTGAACAATTGTGTGGAG	60						
BZIP1	Medtr7g057160	PsCam040141	PsBZIP1-1F	GAAAGCGTCGTCGAGGAAG	58.6	1	147	370 bp	Yes	Yes	Yes
			PsBZIP1-1R	CTTCCATCTCTTTAGCTCTTGATTC	57.5						
CAT9	Medtr7g007850	PsCam026869	PsCAT9_1F	GTCGTCTTCTTTGCATACG	54.3	3	552	No			
			PsCAT9_1R	AGCTGAAGACGATACTTGG	54.3						
			PsCAT9_2F	GCTGGGCTGCTTAATGTGC	59.9	1	143	700 bp	Yes	No	
			PsCAT9_2R	TCTCGCCTAGCTGAAGACG	58.9						
DE1	Medtr7g451010	PsCam050323	PsDE1-1F	GTCAGGTCAGACAAGCAAAACAG	60.2	1	139	300 bp	Yes	Yes	Yes
			PsDE1-1R	AGCCTTTTCGCACACTGTAG	58.5						
DEAD1	Medtr7g057250	PsCam045248	PsDEAD1-1F	CGTTGAATGATGAGAGCATTAAGTC	58.6	1	176	650 bp	Yes	No	
			PsDEAD1-1R	AAGCCTCAACTCGTTCCATG	58.2						
DSP1	Medtr7g055923	PsCam002046	PsDSP1-1F	AGTAAGAGAAACAACGAATCCTGG	58.8	1	141	400 bp	Yes	Yes	Yes
			PsDSP1-1R	AGCTTTACAACCCATGCACAC	59.4						
EAC1	Medtr7g053290	PsCam023290	PsEAC1-1F	CGTAACGAAGGCAGTGTCTG	58.9	1	134	300 bp	Yes	No	
			PsEAC1-1R	GAGATTTTGCTATTGCCCAGC	57.9						
ENTH1	Medtr7g006280	PsCam048827	PsENTH1_1F	ACTCCAGTGGAAGAGATCAAGG	59.2	1	173	600 bp	Yes	No	
			PsENTH1_1R	GCTTGTGTGTGAACCAGGTC	59.3						
EOD1	Medtr7g023770	PsCam007049	PsEOD1-1F	AAGTGGCTCCGTCAGACTG	59.3	1	171	1000 bp	Yes	No	
			PsEOD1-1R	ACACTCCTGAATCTGTTTCCTAAG	58.4						
EZA1	Medtr7g055660	PsCam029108	PsEZA1-1F	GAGGAACACGATTTTACTGAGG	56.8	1	160	No			

Table A1.2. (continued)

Locus name	*M. truncatula* accession	*P. sativum* accession	Primer name	Primer sequence	Tm (° C)	Intron	cDNA size (bp)	PCR product	Sequ-enced	Polym-orphic	Map-ped
			PsEZA1-1R	TGTCCAAACATATTCCCGTAATAG	56.6						
			PsEZA1-2F	AGCAATCAGTTTGGAATGAACC	57.5	1	160	350 bp	Yes	Yes	Yes
			PsEZA1-2R	GCAGATCTGAGACCTTAACATC	56.3						
FADO1	Medtr7g051240	PsCam012946	PsFADO1-1F	TGGACTGGTTCTTTGATGCAG	58.5	1	139	750 bp	Yes	No	
			PsFADO1-1R	CTGCCTCCATCAAGCCATG	58.6						
GOT1	Medtr7g006340	PsCam043494	PsGOT1_1F	TTTTGGCATCTTCTTCACG	53.9	5	374	No			
			PsGOT1_1R	GCGATAGACTTGGAAGCG	55.7						
GPI1	Medtr7g056117	PsCam049521	PsGPI1-1F	GCTGATGGGAGGATTGTTCTTTC	59.6	1	139	700 bp	Yes	Yes	Yes
			PSGPI1-1R	TGACATCGGCTATAGGTTGCAG	60.2						
GS1	Medtr7g005950	PsCam035287	PsGS1_1F	AGAGCTTCGTCTATGGGCATC	59.7	1	110	No			
			PsGS1_1R	TGGCCATATCAAGAAAAGCCTG	59						
GTQ1	Medtr7g057530	PsCam048130	PsGTQ1-1F	CACAGTACATGAATTTCTCGCATC	58.5	1	201	300 bp	Yes	No	
			PsGTQ1-1R	ATGTGATTAGCCCGGGTGG	59.5						
HMB1	Medtr7g005970	PsCam057867	PsHMB1_1F	GCAACATGCGAGATTGAAGC	58.5	1	164	No	No		
			PsHMB1_1R	CCCGATAATCCTCCATTTCTGC	58.9						
HP1	Medtr7g407140	PsCam059449	PsHP1_1F	TTTCCGGCATCTACGTCAAC	58.3	1	152	300 bp	Yes	Yes	Yes
			PsHP1_1R	AGTACAGTTCCGCCTTTGC	58.1						
HPG1	Medtr7g005390	PsCam045353	PsHPG1_1F	TCCAACCACACCAAAACC	55.3	6	588	No			
			PsHPG1_1R	ACTACGTGATTGTCCTTGC	55.2						
ICE1	Medtr7g451440	PsCam045614	PsICE1_1F	TTCCAAGCTCACAGAACACC	58.5	1	131	400 bp	Yes	Yes	Yes
			PsICE1-1R	CCTTTCATCTACCTTGCCCAC	58.6						
L32	Medtr7g053180	PsCam002234	PsL32-1F	AAGAAGCGTGTCAAGAGGTTC	58.5	1	135	350 bp	Yes	Yes	Yes
			PsL32-1R	AATGTTGGGCATCAATGTCAC	57.5	1					
LTP1	Medtr7g405770	PsCam039604	PsLTP1-1F	GTTACCCCTTTGTTCTCATTGC	57.8	1	176	400 bp	Yes	No	
			PsLTP1-1R	GTTGCTTCTTGGTGAGCTCG	59.5						
MAB1	Medtr7g009250	PsCam014151	PsMAB1_1F	TTGATGAGGAAATGTCTGC	52.9	4	317	No			
			PsMAB1_1R	GGGTCCTCTAAAGACAATAGG	54.8						

Table A1.2. (continued)

Locus name	*M. truncatula* accession	*P. sativum* accession	Primer name	Primer sequence	Tm (° C)	Intron	cDNA size (bp)	PCR product	Sequ-enced	Polym-orphic	Map-ped
MCO1	Medtr7g051440	PsCam050297	PsMCO1-1F	TACAAGTTGTAAACCCGGAGAATG	59	1	144	300 bp	Yes	Yes	Yes
			PsMCO1-1R	AATTACTGCTTAAAGGTGATGATCC	57.4						
MGL1	Medtr7g007260	PsCam009698	PsMGL1_1F	GCTACAGAGATAAACCGCGTTTG	60.2	1	146	150 bp	No		
			PsMGL1_1R	TTTGCTCACCGACGGATCC	60.1						
NADP1	Medtr7g406860	PsCam033498	PsNADP1_1F	TGCGATCCCTATATGCGTGG	59.8	1	148	500 bp	Yes	No	
			PsNADP1_1R	CGGAAATGAAATCACCAGCTTTG	59.1						
NEF1	Medtr7g406760	PsCam042442	PsNEF1_1F	TGGGTATGGTTTCTGATGC	54.8	1	210	No			
			PsNEF1_1R	ACCATGTGACCATCAAGC	54.9						
NTF2	Medtr7g055610	PsCam030850	PsNTF2-1F	GTAAGCTTGGTCGTCCTGAAC	58.9	1	140	600 bp	Yes	Yes	Yes
			PsNTF2-1R	ACCATAGGACTCTTGTGCATCC	59.8						
PAP1	Medtr7g055923	PsCam034372	PsPAP1-1F	GCCAGTCGCTAATATACAATCAAC	58.1	1	143	750 bp	Yes	Yes	Yes
			PsPAP1-1R	GGCATGTAATGAAGGATCTCCAC	59.2						
PDE1	Medtr7g005380	PsCam014151	PsPDE1_1F	GGTCTAAAGCCTGTTGTGGAG	58.6	1	104	1200 bp	yes	No	
			PsPDE1_1R	TGCCCAGCAGACATGTAATTTG	59.5						
PPC1	Medtr7g021530	PsCam010674	PsPPC1-1F	GGGTCAAGTGATGTTGCTTCC	59.5	1	150	No			
			PsPPC1-1R	AACCCTATGGCCATAAGGTCC	59.2						
PPR1	Medtr7g056613	PsCam000159	PsPPR1-1F	TATCAAACCAAACACAGTCACCTAC	59.2	1	150	500 bp	Yes	No	
			PsPPR1-1R	GCCCTGAGCGTTGAATTCATG	60.2						
PT1	Medtr7g006540	PsCam029064	PsPT1_1F	AGTCCTCTTACACGTGTCGC	59.8	1	232	450 bp	Yes	No	
			PsPT1_1R	GTGTGACTGTGCAAAGGTTCC	59.9						
PUB9	Medtr7g005940	PsCam011495	PsPUB9_1F	AGATTGGCTCGGTGAAGG	56.6	3	516	No			
			PsPUB9_1R	ATAGCTTTCCGTTCGAAGG	55						
RAD23	Medtr7g407040	PsCam034125	PsRAD23_1F	GGCCAAGCAGAATCTAATCTCC	58.9	1	212	900 bp	Yes	Yes	Yes
			PsRAD23_1R	TCAACATTGGTGGATGCAGC	59.1						
RBC1	Medtr7g007120	PsCam011201	PsRBC1_1F	ATTCGGCGGCCTCAAATCC	60.8	1	162	No			

Table A1.2. (continued)

Locus name	*M. truncatula* accession	*P. sativum* accession	Primer name	Primer sequence	Tm (° C)	Intron	cDNA size (bp)	PCR product	Sequ-enced	Polym-orphic	Map-ped
			PsRBC1_1R	CTCGTCAATGGTGGCAAATAGG	59.6						
S62	Medtr7g053160	PsCam011292	PsS62-1F	TGTCTAAGGAGGATGATGTGAGG	59	1	140	250 bp	Yes	yes	Yes
			PsS62-1R	TCTTGCTCGTTTTCTTTGGAGAG	59.2						
SAG29	Medtr7g405730	PsCam011330	PsSAG29_1F	TCAGTCACAGCACTTTAGC	55.2	3	381	No			
			PsSAG29_1R	CACGAACCGAACCATGC	56.3						
SAMT1	Medtr7g006060	PsCam023112	PsSAMT1_1F	GTTTGTGTTTCCTGGTGG	53.8	2	415	No			
			PsSAMT1_1R	TACCCAAAATCCACCAGG	53.6						
SCF35	Medtr7g058520	PsCam045631	PsSCF35_1F	GTTTGTTCATATGCCCCTTTTGG	58.7	1	162	650 bp	Yes	No	
			PsSCF35_1R	GCTTTAACCATGTAAACCTCTCAC	58.1						
SPB1	Medtr7g028740	PsCam012994	PsSPB1-1F	GGAAACCTCAACCACCTTCTC	58.5	1	149	550 bp	Yes	No	
			PsSPB1-1R	GGCAGGCCAAATGTTTCTC	57.2						
SPP1	Medtr7g405780	PsCam023099	PsSPP1-1F	TCCTGAGAAGTCCGGTGAAC	59	1	192	250 bp	yes	No	
			PsSPP1-1R	AGCACCCACATAAACCATGC	58.8						
SSP1	Medtr7g007010	PsCam057865	PsSSP1_1F	AGCTAAATTCTACGGTCGTGC	58.5	1	164	300 bp	yes	No	
			PsSSP1_1R	AGCTCCCCGACCATAGTAAAC	59.2						
SWEET9	Medtr7g007490	PsCam039934	PsSWEET9_1F	ACATTTTCGGATCATGAGATGG	56.6	3	410	No			
			PsSWEET9_1R	GCACAAATCAAACCAACAGC	56.7						
TDF1	Medtr7g006560	PsCam044291	PsTDF1_1F	AAGGCTCCTTTGGGCTTCTAC	60	1	178	300 bp	Yes	Yes	Yes
			PsTDF1_1R	TCCTTCAGCTTCAGATGCATTTG	59.6						
THP1	Medtr7g009070	PsCam010122	PsTHP1_1F	AACCGATTGTTCATGCCATGG	59.5	1	161	400 bp	Yes	Yes	Yes
			PsTHP1_1R	CGCACCAACAATCTTATCCACC	59.9						
UPF1	Medtr7g052250	PsCam038395	PsUPF1-1F	CATGGGCAAGCCATCATTTAATG	58.9	1	134	550 bp	Yes	No	
			PsUPF1-1R	TTTTATGCTCTGGCTTTGTGTC	57.5						
VAS1	Medtr7g007770	PsCam040691	PsVAS1_1F	TCAAAGGTCAAGGTGGCATTC	58.8	1	135	300 bp	Yes	Yes	No
			PsVAS1_1R	ATTCAGCTGCCTCTTTCTCAG	58						

Table A1.3. Details on all other primers that were used for various other purposes involving pea apart from finding polymorphism between parental genotypes which are mentioned in table A1.1 and A1.2

Locus name	*M. truncatula* accession number	*P. sativum* accession number	Pea *LG*	Primer name	Primer sequence	Tm (° C)	Purpose	Reference
AAP1	Medtr3g088475	PsCam048093	*LGIII*	AAP1-F	GTCATGATCCTCTTTGCTTGGG	55.9	CAPS marker	Aubert et al., 2006
			LGIII	AAP1-R	CCAAGATTCGTATGAACAACTTCC	54	CAPS marker	Aubert et al., 2006
AAP2	Medtr3g096830	PsCam049915	*LGIII*	AAP2-F	TTTGGACCATATATGGCATATGC	53.1	Size marker	Aubert et al., 2006
			LGIII	AAP2-R	CAATAAAATGCAGCAATCACAGCC	55.6	Size marker	Aubert et al., 2006
AAR3	Medtr7g010920	PsCam050275	*LGV*	MtAAR3-1F	CTTCATGTGTCGTGAAAATGG	52.2	CAPS marker	Hecht et al., unpublished data
			LGV	MtAAR3-1R	GAACCTGTTGCCACGTATCC	56	CAPS marker	Hecht *et. al.*, unpublished data
ABHP1	Medtr7g007310	PsCam038878	*LGV*	ABHP1-TER-HRM-F	CCGGCAACTGAATCAACCG	59.5	HRM marker	This study
			LGV	ABHP1-TER-HRM-R	TGCTACGCTCTCTCCTTGC	59.5	HRM marker	This study
ABI3	Medtr7g059330	PsCam027818	*LGV*	ABI3-F	GTGTATCAGCTAATGATTTGAGG	51.4	CAPS marker	Aubert et al., 2006
			LGV	ABI3-R	CTGCCAATACATCCAGTTAGC	53.5	CAPS marker	Aubert et al., 2006
ATPB1	Medtr7g055720	PsCam033465	*LGV*	ATPB1-HRM-1F	TTTCCCCAAATTTATTGTTACAAGC	57	HRM marker	This study
			LGV	ATPB1-HRM-1R	CCTTTGGTCACTGTTTGTTATCAAC	59	HRM marker	This study
ATX6	Medtr3g095840	PsCam038017	*LGIII*	PsATX6-1F	GTTCACGCAAAGGAGAATACTTG	58.3	Sequence candidate gene	This study
			LGIII	PsATX6-2R	AAAGACCCGCGATTCATCATTAAC	60	dCAPS marker	This study
			LGIII	PsATX6-3R	GTGGACTGATTCGCTTCTTG	56.8	Sequence candidate gene	This study
			LGIII	PsATX6-3F	CCCATTTTCTAGAGCTTACACC	56.5	Sequence candidate gene	This study
			LGIII	PsATX6-3R2	CGAGAAGCAGAATTATTCCGTC	57.3	Sequence candidate gene	This study
			LGIII	PsATX6-4F	GAAGAAAAGGAGAAGGTTACTGC	57.4	Sequence candidate gene	This study
			LGIII	PsATX6-4R	TTGAATCTGCAAGGACTATAAGC	56.7	Sequence candidate gene	This study
			LGIII	ATX6-HindII-F	AAATACTTATACTTTTTCTTTAGTAAGTTGA	50.4	dCAPS marker	This study

Table A1.3. (continued)

Locus name	*M. truncatula* accession number	*P. sativum* accession number	Pea *LG*	Primer name	Primer sequence	Tm (° C)	Purpose	Reference
BTB1	Medtr3g096160	PsCam000088	*LGIII*	BTB1-TER-F2	TCCATCATTTTATGGAATGCAA	54.4	HRM marker	This study
			LGIII	BTB1-TER-R2	GGTAAAAACAAGAGAAAACCAA	53.3	HRM marker	This study
BZIP1	Medtr7g057160	PsCam040141	*LGV*	BZIP1-HRM-1F	TGTTCCTTAGCAACCTGTATG	55.6	HRM marker	This study
			LGV	BZIP1-HRM-2R	AAGATCAGATGGAGTATAAGACAAG	56	HRM marker	This study
CDKE1/CDK8	Medtr3g096960	PsCam048317	*LGIII*	PsCDKE1-1F	TGCAGAGCAGTTAACCGATTG	58.9	Sequence candidate gene	This study
			LGIII	PsCDKE1-1R	TCGATGGCTTCAAATCTCGATG	58.9	Sequence candidate gene	This study
			LGIII	PsCDKE1-2F	GACATCATAGAGACAAGGTCAACC	58.9	Sequence candidate gene	This study
			LGIII	PsCDKE1-2R	GCTGCTGTTATACGCTTTTTAGG	58.4	Sequence candidate gene	This study
			LGIII	PsCDKE1-3F	TGCTAATCTCTACAGTGTTGTACAC	58.6	Sequence candidate gene	This study
			LGIII	PsCDKE1-3R	AAGCCAAGTTTCTCTATGACAGAG	58.5	Sequence candidate gene	This study
			LGIII	PsCDKE1-4F	ACATCGAGATTTGAAGCCATCG	59.1	Sequence candidate gene	This study
			LGIII	PsCDKE1-4R	CCTTGCAAGCCCAAAGTCAG	59.7	Sequence candidate gene	This study
			LGIII	PsCDKE1-5F	TTCAAAATGGAACCACAGCCTG	59.6	Alternative splicing in *late3-2*	This study
			LGIII	PsCDKE1-5R	GACCACCAGGCATGTTTCC	58.7	Alternative splicing in *late3-2/late3-3*	This study
			LGIII	PsCDKE1-6F	AAGCCATCGAATATATTGATACGAG	57	Alternative splicing in *late3-3*	This study
			LGIII	CDKE1-HRM-1F	TCTGATTCGTGGTCGCGTC	60.2	HRM marker	This study
			LGIII	CDKE1-HRM-1R	TCAGCCATGGTTGCTCCTC	59.7	HRM marker	This study
			LGIII	CDK8-Y2H-F	ACCATGGCTGACGGCAACC	63.2	Yeast two hybrid experiment	This study
			LGIII	CDK8-Y2H-R	TCACATACGTCTTGGCTTCTGC	60.9	Yeast two hybrid experiment	This study

Table A1.3. (continued)

Locus name	*M. truncatula* accession number	*P. sativum* accession number	Pea *LG*	Primer name	Primer sequence	Tm (° C)	Purpose	Reference
CYCC1	Medtr7g055650	PsCam050605	*LGV*	PsCYCC1-1F *	ACCAAGAACAACACACGAACTAG	59.1	Sequence candidate gene	This study
			LGV	PsCYCC1-2F *	AGGCTAGGCTGCTTGTATTTTAC	58.8	Sequence candidate gene/HRM marker	This study
			LGV	PsCYCC1-2R *	ACGCACCAAGCCAAACTC	58.6	Sequence candidate gene	This study
			LGV	PsCYCC1-3F	AGCAGTTACTTATATGAGGCGTG	58.4	Sequence candidate gene/alternative splcing in *late4-1/late4-2*	This study
			LGV	PsCYCC1-3R	TGCCTCACTTTTACTTGTTGAGC	59.7	Sequence candidate gene	This study
			LGV	PsCYCC1-4R	TACAAGCAGGTTGGAGCTACC	59.7	Sequence candidate gene	This study
			LGV	PsCYCC1-5F	TGACACATACAAGATGGACCTAATG	58.7	Sequence candidate gene	This study
			LGV	PsCYCC1-5R	TATCAACTCGAAGTTCTTCATACC	56.3	Sequence candidate gene/alternative splcing in *late4-1/late4-2*	This study
			LGV	PsCYCC1-7F	GGATCAAGAAGATGTGGATATGGTG	59.3	Sequence candidate gene	This study
			LGV	PsCYCC1-8F	TCTCGTGTTTGGACTTTAGGAATTC	59.3	Sequence candidate gene	This study
			LGV	PsCYCC1-8R	AACACGCCTCATATAAGTAACTGC	59.2	Sequence candidate gene	This study
			LGV	PsCYCC1-9R	GTGGAGAGCAGCATTGATTCTC	59.4	Sequence candidate gene	This study
			LGV	PsCYCC1-9F	CATCGTGGCACGCCTAATATG	59.5	Sequence candidate gene	This study
			LGV	PsCYCC1-7R	AAGACGACTTTATTGGCTAACTAAC	57.4	HRM marker	This study
			LGV	CYCC1-HRM-1F	GTATCTGGCATCAAAAGCAGAAG	58.1	Genotype *late3-1* x *late 4-2*_F3	This study
			LGV	CYCC1-HRM-1R	AATGTAAAATACAAGCAGCCTAGC	57.9	Genotype *late3-1* x *late 4-2*_F3	This study
			LGV	CYCC1-Y2H-1F	ATGGCTGCCAATTTCTGG	59.6	Yeast two hybrid experiment	This study
			LGV	CYCC1-Y2H-1R	AAGATCATAGCTTGTGGAGAGC	57.9	Yeast two hybrid experiment	This study
DE1	Medtr7g451010	PsCam050323	*LGV*	DE1-HRM-1F	CATCTATCTCAAATGTTGGACGAAG	58.1	HRM marker	This study
			LGV	PsDE1-1R	AGCCTTTTCGCACACTGTAG	58.5	HRM marker	This study
DET	Medtr7g104460	PsCam040094	*LGV*	TFL1a-1F	CGTTGGTAGAGTCATAGG	58	Q-RT-PCR	Hecht et al., 2011
			LGV	TFL1a-2R	AGGATCACTAGGGCCAGG	58	Q-RT-PCR	Hecht et al., 2011
DSP1	Medtr7g055923	PsCam002046	*LGV*	DSP1-HRM-1F	TCCCTCATACATACCGTAAATAGC	57.7	Size marker	This study
			LGV	DSP1-HRM-1R	CACGTACACATACACATACACATAC	57.9	Size marker	This study

*Primers marked in red color, *PsCYCC1* = *PsCDKE2* (within the primer naming system of the group of Dr. Jim Weller)

Table A1.3. (continued)

Locus name	*M. truncatula* accession number	*P. sativum* accession number	Pea *LG*	Primer name	Primer sequence	Tm (° C)	Purpose	Reference
ELA1	Medtr3g093530	PsCam000417	*LGV*	ELA1-SNP-1F	GCTTGTTTTGGGAGCAATTATG	56.8	Sequence candidate gene	This study
			LGV	ELA1-SNP-1R	AGCTAATAACATTAAGCACCACG	57.3	Sequence candidate gene	This study
			LGV	PsELA1-7F2	CGAAGCAACGGTAGGTAGTG	58.4	Sequence candidate gene	This study
			LGV	PsELA1-7R	TGCTGAGTGAGGAGCATAATC	57.5	Sequence candidate gene	This study
			LGV	ELA1-HRM-1F	TCCCAGAGTTCGGAGGAAAG	58.7	HRM marker	This study
			LGV	ELA1-HRM-1R	ACCATTAACCATTGTATACTCCCAG	58.4	HRM marker	This study
EZA1	Medtr7g055660	PsCam029108	*LGV*	EZA1-cDNA-1F	TGCTTCCTTTGCTCGGTTAAC	59.1	Sequence candidate gene	This study
			LGV	EZA1-cDNA-1R	ACAACACTCAATGCTTCCTCAG	58.9	Sequence candidate gene	This study
			LGV	EZA1-cDNA-2F	TCTCAGGGTGAAGACCAAATTC	58	Sequence candidate gene	This study
			LGV	EZA1-cDNA-2R	GCATCTCATCAACTTGCCATTC	58.3	Sequence candidate gene	This study
			LGV	EZA1-cDNA-3F	GCTGAAGGGAATCCTAGAGAAC	57.9	Sequence candidate gene	This study
			LGV	EZA1-cDNA-3R	CTCCTAATGAACCATCGCCAC	58.8	Sequence candidate gene	This study
			LGV	EZA1-cDNA-4F	ATAGGCCAATGCAGAAATGGAC	59	Sequence candidate gene	This study
			LGV	EZA1-cDNA-4R	CACACTCCCCTTTATCCATAGC	58.3	Sequence candidate gene	This study
			LGV	EZA1-HRM-1F	TTGCTTGGTGTTTCAACTTCTATC	57.9	HRM marker	This study
			LGV	EZA1-HRM-1R	TGAAGTATTCACATCCTTTATCTGC	57.2	HRM marker	This study
FRI	Medtr3g098290	PsCam036412	*LGIII*	FRI-F3	TGCAACCATTTGTTTTAAGGTC	56.6	Size marker	Hecht et al., unpublished data
			LGIII	FRI-R3	AGGGAAATTTTGGGTGGAAT	54.3	Size marker	Hecht et al., unpublished data
FTa1	Medtr7g084970	PsCam034177	*LGV*	FTLa-6F	GCCCAAGCAACCCTACTTTT	60	Q-RT-PCR	Hecht et al., 2011
			LGV	FTLa-2R	CCATCCTGGAGCGTAAACCC	60	Q-RT-PCR	Hecht et al., 2011
FTb2	Medtr7g006690	PsCam024839	*LGV*	FTLe2-F7	CGACTACCGGGACAGCATTT	62	Q-RT-PCR	Hecht et al., 2011
			LGV	FTLe2-R7	CAGGTGAACCAAGGTTATAAAC	62	Q-RT-PCR	Hecht et al., 2011
FTc	Medtr7g085040	PsCam040405	*LGV*	FTLc-8F	GATATTCCAGCCACAACAAGC	62	Q-RT-PCR	Hecht et al., 2011
			LGV	FTLc-7R	TTATGACGCCACTCTGGAGCAA	62	Q-RT-PCR	Hecht et al., 2011
GPI1	Medtr7g056117	PsCam049521	*LGV*	GP1-HRM-1F	CAGTTTAAACTCCACTGGCTG	57.1	Size marker	This study
			LGV	GP1-HRM-1R	ACAATATTCAGAGCGATAACTAAGC	57.4	Size marker	This study
HP1	Medtr7g407140	PsCam059449	*LGV*	PsHP1_1F	TTTCCGGCATCTACGTCAAC	58.3	CAPS marker	This study
			LGV	PsHP1_1R	AGTACAGTTCCGCCTTTGC	58.1	CAPS marker	This study

Table A1.3. (continued)

Locus name	*M. truncatula* accession number	*P. sativum* accession number	Pea *LG*	Primer name	Primer sequence	Tm (° C)	Purpose	Reference
ICE1	Medtr7g451440	PsCam045614	*LGV*	ICE1-HRM-1F	ACGGTGTCCTTGATATATGAGTATG	58.1	HRM marker	This study
			LGV	ICE1-HRM-1R	CGCTCGTTTTGCTTCCTATTTG	58.8	HRM marker	This study
L32	Medtr7g053180	PsCam002234	*LGV*	L32-HRM-1F	GGAGAATATGTGTCAAGGTATGTC	56.9	HRM marker	This study
			LGV	L32-HRM-1R	CATTTGACACAACATAGCACAAAAC	58.4	HRM marker	This study
LCV2	Medtr3g095210	PsCam021197	*LGIII*	LCV2-HRM-F	TGAGTAAATTGAGAGAAAAGGC	54.4	HRM marker	This study
			LGIII	LCV2-HRM-R	ACACACACGTTTTGAAGC	54.2	HRM marker	This study
LF	Medtr1g060190	PsCam040094	*LGII*	LF-CR2	AAATAAGCAGCAGCAACAGGG	60	Q-RT-PCR	Foucher et al., 2003
			LGII	LF-CR3	CAGACATTCCAGGGACAACAG	60	Q-RT-PCR	Foucher et al., 2003
MCO1	Medtr7g051440	PsCam050297	*LGV*	MCO1-HRM-1F	CAAAGGTAAGATCGCCTTAAAAC	56.2	HRM marker	This study
			LGV	PsMCO1-1R	AATTACTGCTTAAAGGTGATGATCC	57.4	HRM marker	This study
NIP2	Medtr3g099380	PsCam013125	*LGIII*	NIP-F	GGATTCTACTAACCACATGGACC	54.8	CAPS marker	Aubert et al., 2006
			LGIII	NIP-R	CTGCATGTCAGAGGTCCATAGG	57.1	CAPS marker	Aubert et al., 2006
NTF2	Medtr7g055610	PsCam030850	*LGV*	NTF2-HRM-1F	CACATGGGAAGTGAAGTAAAAGC	58.2	HRM marker	This study
			LGV	NTF2-HRM-1R	ACATCATGTGTACCCCTGTTTC	58.3	HRM marker	This study
PAP1	Medtr7g055923	PsCam034372	*LGV*	PAP1-HRM-1F	AGTTGTAACTGTATGAAACTTTCAC	56	HRM marker	This study
			LGV	PAP1-HRM-1R	TAATAGTATAGTGTGCTTGGTTCC	56	HRM marker	This study
PIM	Medtr8g066260	PsCam046356	*LGIV*	PIM-4F	GCTTCAGAGTTTGGAACAGC	58	Q-RT-PCR	Hecht et al., 2011
			LGIV	PIM-6R	GACTCCATGGTGGTTTGG	58	Q-RT-PCR	Hecht et al., 2011
PM10 /VEG1	Medtr7g016630	PsCam001681	*LGV*	FULC-2F	CGATGCCTTGAAACCATAGG	53.6	Size marker	Hecht et al., unpublished data
			LGV	FULC-3R	AGATGTTGTAGCATCCATGC	52.8	Size marker	Hecht et al., unpublished data
			LGV	FULC-2F	CGATGCCTTGAAACCATAGG	58	Q-RT-PCR	Hecht et al., 2011
			LGV	FULC-2R	AATTCCAATGACCCTCTTGC	58	Q-RT-PCR	Hecht et al., 2011
RAD23	Medtr7g407040	PsCam034125	*LGV*	RAD23-ApoI_F	CCAGTTCTGCTATCTTGGTTCATGAATT	57.3	dCAPS marker	This study
			LGV	RAD23-ApoI_R	AAATCGCTTCCACTAGGGCTG	57.3	dCAPS marker	This study
RBCS	Medtr7g007230	PsCam011201	*LGV*	RBCS-ASP1	TGCTTCTACGGTGCAATCG	53	ASP marker	Aubert et al., 2006
			LGV	RBCS-ASP2	GCTCACGGTACACAAATCC	53	ASP marker	Aubert et al., 2006
			LGV	RBCS-ASP3	CCATTAATACTATTCAAATAGC	53	ASP marker	Aubert et al., 2006
			LGV	RBCS-ASP4	ATTTGAGAACTAGAAGACTATA	53	ASP marker	Aubert et al., 2006

Table A1.3. (continued)

Locus name	*M. truncatula* accession number	*P. sativum* accession number	Pea *LG*	Primer name	Primer sequence	Tm (° C)	Purpose	Reference
S62	Medtr7g053160	PsCam011292	*LGV*	S62-Taq1-1F	ATCACACACATACAAATTCTACTAAAAATC	52.8	dCAPS marker	This study
			LGV	S62-HRM-1R	AACAAAAGCTGGTAAGATCTTGG	57.3	dCAPS marker	This study
TDF1	Medtr7g006560	PsCam044291	*LGV*	TDF1-HRM-F	TTTTCCAATGTCCATTTGCTTATAG	56.4	HRM marker	This study
			LGV	TDF1-HRM-R	TACAGCACCATGATCGATAAATG	56.8	HRM marker	This study
TFIIa	Medtr2g097740	PsCam049917		TFIIa-qF	GCAACCTCCTTCTCCTTGGAT	60	Q-RT-PCR	Hecht et al., unpublished data
				TFIIa-qR	TCTTCCCGTCCTTCCACATAA	60	Q-RT-PCR	
THP1	Medtr7g009070	PsCam010122	*LGV*	THP1-HRM-F	GCCATGGCTAACGAATTC	53.9	HRM marker	This study
			LGV	THP1-HRM-R	GTCAACGTCGATCTTGATG	54	HRM marker	This study
TPR1	Medtr3g095070	PsCam038860	*LGIII*	TPR1-TER-HRM-F	TGGTGTGAAATTGAATGCTTTG	56.3	HRM marker	This study
			LGIII	TPR1-TER-HRM-R	ATGTCAATCCAAACACACCT	55.1	HRM marker	This study
UNI	Medtr3g098560	PsCam056640	*LGIII*	UNI-1F	CATCAGAGCTGAAAGAAGG	50.4	CAPS marker	Hecht et al., unpublished data
			LGIII	UNI-2R	GCTTCCTTTTCACGTTGC	52.3	CAPS marker	Hecht et al., unpublished data
			LGIII	UNI-1F	CATCAGAGCTGAAAGAAGG	55	Q-RT-PCR	Hecht et al., 2011
			LGIII	UNI-2R	GCTTCCTTTTCACGTTGC	55	Q-RT-PCR	Hecht et al., 2011
VEG2	Medtr5g022780	PsCam027284	*LGI*	FD-6F	ATTTGATCCAAACGTCGGTGT	60	Q-RT-PCR	Hecht et al., 2011
			LGI	FD-3R	ATCAACTTTTTGCTCCAGTTCG	60	Q-RT-PCR	Hecht et al., 2011
WRI11	Medtr7g009410	PsCam038324	*LGV*	WRI1-1F	GATTTCGTGGTGTTAGCAGG	53.7	CAPS marker	Hecht et al., unpublished data
			LGV	WRI1-2F	CGATTTGAAGCTCACTTGTGG	54.3	CAPS marker	Hecht et al., unpublished data
				PGEMT-F	GCCCGACGTCGCATGCTCC	55	Colony PCR for pea CYCC1	Hecht et al., unpublished data
				PGEMT-R	GAGCTCTCCCATATGGTCG	55	Colony PCR for pea CYCC1	Hecht et al., unpublished data

Table A1.3. (continued)

Locus name	***M. truncatula* accession number**	***P. sativum* accession number**	**Pea LG**	**Primer name**	**Primer sequence**	**Tm (° C)**	**Purpose**	**Reference**
				TOPO-1F	CGCTAGCATGGATGTTTTCC	55	Colony PCR during Y2H experiment	Hecht et al., unpublished data
				pDEST32-F	CGCTACTCTCCCAAAACCAA	60	Confirmation of LR reaction during Y2H	Hecht et al., unpublished data
				pDEST22-F	TGAAGATACCCCACCAAACC	60	Confirmation of LR reaction during Y2H	Hecht et al., unpublished data

Table A1.4. Primers designed in *A. thaliana* which were used in this study

Locus name	Accession number	SALK-ID	Primer position	Name	Tm(° C)	Purpose	Reference
AtCYCC1_1	AT5G48640	SALK_039400	AtCYCC1-LP2	TAATGTGTTGTGCCAGTTTCG	59.6	Gene specific PCR	This study
			AtCYCC1-RP2	CTATTTTCCGTGTTTCAGGGG	60.7	Gene specific PCR/Insertion specifc PCR	This study
AtCYCC1_2	AT5G48640	SALK_053291	AtCYCC1-LP1	AGAACCGTTGTCGTCGATATG	60.0	Gene specific PCR	This study
			AtCYCC1-RP1	TGCCAATTTCTGGAATTCATC	59.9	Gene specific PCR/Insertion specifc PCR	This study
-	-	-	AtLBb1.3	ATTTTGCCGATTTCGGAAC	60.0	Insertion specific PCR	This study

Appendix 2 -Tables, sequence alignments and figures for chapter 5

Table A2.1. Comparison of low and high coverage RNA sequencing data analysis for pea transcriptome sequences located between markers *BTB1-SPS1* in *PsLGIII*

P. sativum accession number	Low coverage run				High coverage run			
	WT (NGB5839)		*late3-1*		WT (*late5*)		*late3-1*	
	Read number	Ref sequence (%)	Read number	Ref sequence (%)	Read number	Ref sequence (%)	Read number	Ref sequence (%)
PsCam000088	80	90	44	66.1	435	97.6	537	94.2
PsCam034774	109	79.8	73	71.9	602	95.3	819	79.3
PsCam035309	50	25.3	36	24.6	207	49.4	277	58.1
PsCam048280	145	81	79	70.2	569	95.9	798	96.7
PsCam000112	4	12	12	31.6	53	78.1	75	83.0
PsCam001940	168	83.5	42	71.8	878	96.3	498	96.5
PsCam034645	20	67.6	14	36.7	68	82.8	100	81.3
PsCam030881	42	93.5	6	20.2	103	100.0	148	99.5
PsCam052035	35	80.7	11	46.2	130	87.2	138	94.3
PsCam037303	11	40.2	5	16.5	49	81.6	70	82.4
PsCam029411	30	65.8	21	61.1	165	78.2	194	78.9
PsCam050053	4	12.8	6	12.9	32	48.9	51	63.6
PsCam059372	-	-	2	5.7	2	7.1	2	6.6
PsCam049851	20	56.7	13	29.5	109	92.4	152	86.0
PsCam039877	-	-	-	-	2	10.9	1	2.6
PsCam032493	496	100	289	98.6	2826	100.0	2915	100.0
PsCam046205	8	32.3	5	16.3	96	91.2	20	45.0
PsCam004937	14	48.5	2	5.3	52	77.2	53	65.6
PsCam001809	266	99.9	151	87.4	1358	100.0	1923	100.0
PsCam008813	66	77.1	49	68.4	418	95.0	439	94.7
PsCam033826	308	84.2	120	75	1051	91.7	1113	91.2
PsCam001401	4	25.7	-		19	52.5	18	60.5
PsCam054769	22	60.6	20	42.8	111	88.8	109	78.8
PsCam034500	49	61	24	44.4	204	94.6	205	87.2
PsCam037090	27	52.5	19	48.4	157	94.0	212	94.0
PsCam046150	17	56.7	9	28.9	66	73.0	114	85.7
PsCam000535	27	58.3	14	45.4	231	98.8	279	99.8
PsCam039482	15	59.7	4	22.9	69	96.1	66	91.3
PsCam012934	7	17.7	-	-	3	9.0	18	15.4
PsCam050308	89	86.6	58	65.3	363	92.4	506	95.2
PsCam042463	21	26.1	18	27.3	142	70.7	202	79.8
PsCam049915	80	80.6	24	40	398	98.7	252	93.4
PsCam021158	806	92.3	554	92	5784	97.7	5306	99.2
PsCam049087	251	91.6	156	82.6	1286	96.7	1580	97.3
PsCam039677	28	88.8	4	16.5	227	87.3	62	76.9
PsCam042665	121	87.4	80	78.7	591	91.1	725	92.4
PsCam048115	233	96.5	108	85.5	897	100.0	1447	98.0
PsCam048369	24	27.2	10	15.3	141	68.3	160	61.0
PsCam037556	10	47.2	7	35.1	31	63.6	45	61.4
PsCam048317	64	73.3	26	41.3	346	97.3	253	91.6
PsCam035372	5	16	10	33.8	60	58.3	82	78.7
PsCam035385	13	35.9	12	29	48	79.4	93	76.7
PsCam042743	73	83.4	43	73.2	383	96.9	522	90.5
PsCam011333	55	98.2	46	87.1	331	100.0	408	100.0

Table A2.1 (continued)

P. sativum accession number	Low coverage run				High coverage run			
	WT (NGB5839)		*late3-1*		WT (*late5*)		*late3-1*	
	Read number	Ref sequence (%)	Read number	Ref sequence (%)	Read number	Ref sequence (%)	Read number	Ref sequence (%)
PsCam023312	13	40.8	3	8.9	73	90.1	80	79.9
PsCam034443	38	60.8	38	47.5	271	96.3	332	97.5
PsCam039622	21	59.1	43	76	169	83.9	262	78.0
PsCam010752	61	90	36	76.2	271	99.9	342	100.0
PsCam042642	69	89.9	32	50	257	98.0	323	95.8
PsCam034424	32	45.3	30	48.7	211	84.2	324	85.7
PsCam059306	-	-	1	3	6	9.7	2	3.1
PsCam037057	56	80.3	32	58.3	257	90.1	270	90.3
PsCam034607	98	79.4	45	55.1	606	95.3	476	87.8
PsCam050549	107	91.9	65	68.9	689	97.5	671	93.8
PsCam036496	115	76.4	87	74.8	599	88.0	854	86.4
PsCam001579	5	29.2	-	-	20	42.6	12	27.0
PsCam042403	342	97.8	207	89.7	1499	99.2	1925	99.9
PsCam049636	98	83.7	65	88.7	483	100.0	693	100.0
PsCam035950	101	88.4	59	64.8	508	100.0	622	100.0
PsCam033289	114	90.3	74	73.1	577	90.7	776	90.7
PsCam046255	9	47.1	2	11.1	25	79.0	45	86.6
PsCam045416	152	94.8	114	83.9	717	98.8	1224	95.2

Table A2.2. Low coverage RNA sequencing data analysis for pea transcriptome sequences located between markers *S62-OH1* in *PsLGV*

P. sativum accession number	WT		*late4-1*	
	Number of reads	Ref sequence (%)	Number of reads	Ref sequence (%)
PsCam011292	10,497	100	5419	92.3
PsCam002234	994	97.9	983	98.6
PsCam038370	-	-	3	6
PsCam021303	31	80	19	52.1
PsCam026961	155	99.6	112	95.1
PsCam035812	6	13.9	4	7.8
PsCam050024	42	55.4	34	34.2
PsCam023290	164	97.7	132	98.7
PsCam039273	-	-	-	-
PsCam042823	463	94.9	280	98
PsCam037771	-	-	2	5.5
PsCam042714	22	28.5	33	56.5
PsCam021339	2253	87.3	1870	93.7
PsCam043919	587	97.8	887	86.3
PsCam050614	255	93.6	240	93.7
PsCam012869	47	79.8	30	69.3
PsCam045678	114	98.7	71	86.6

Table A2.2. (continued)

P. sativum accession number	WT		*late4-1*	
	Number of reads	Ref sequence (%)	Number of reads	Ref sequence (%)
PsCam001210	42	78.3	30	67.7
PsCam035395	88	78.9	38	68.5
PsCam000439	54	66.9	32	41
PsCam048130	148	91.7	106	82.6
PsCam030850	218	94.7	185	93
PsCam036022	16	12.8	16	22.1
PsCam029431	14	61.6	2	8.6
PsCam035904	77	83.1	70	68.3
PsCam050605	46	86.3	6	27.4
PsCam029108	1	3.4	-	-
PsCam037210	47	100	21	54.2
PsCam024790	107	83.3	80	90
PsCam033465	15	44.1	15	43
PsCam017510	43	68.8	27	59.3
PsCam000368	9	30.4	2	4.5
PsCam010827	14	46.2	11	29.9
PsCam043857	135	70.3	108	84.3
PsCam029356	19	64.7	10	31.2
PsCam007040	35	73.4	30	66.5
PsCam052326	-	-	-	-
PsCam002046	10	53.8	19	68.6
PsCam045822	20	50.2	18	49.2
PsCam056096	11	37	11	48.1
PsCam038127	1	5	3	7.9
PsCam055630	-	-	-	-
PsCam033914	-	-	6	16
PsCam049521	55	70.9	46	71.9
PsCam038311	6	31	4	18.2
PsCam036246	65	84.5	42	72
PsCam045481	78	78.3	48	60.6
PsCam023210	7	35.4	12	43.7
PsCam056707	-	-	1	9.8
PsCam049679	46	68.9	32	55.8
PsCam042433	77	62.6	90	71
PsCam014058	104	80.1	62	76.6
PsCam050951	76	86	79	92.4

Table A2.3. Name and accession numbers of *A. thaliana* Cyclin dependent kinases.

Name	Accession number	Group	Source
AtCDKA.1	AT3G48750	Cyclin dependent kinase A	Tank et al. 2009
AtCDKB1.1	AT3G54180	Cyclin dependent kinase B	
AtCDKB1.2	AT2G38620		
AtCDKB2.1	AT1G76540		
AtCDKB2.2	AT1G20930		
AtCDKC.1	AT5G10270	Cyclin dependent kinase C	
AtCDKC.2	AT5G64960		
AtCDKD1.1	AT1G73690	Cyclin dependent kinase D	
AtCDKD1.2	AT1G66750		
AtCDKD1.3	AT1G18040		
AtCDKE.1/AtCDK8	AT5G63610	Cyclin dependent kinase E	
AtCDKF.1	AT4G28980	Cyclin dependent kinase F	
AtCDKG.1	AT5G63370	Cyclin dependent kinase G	
AtCDKG.2	AT1G67580		
AtCDKL.1	AT5G39420	Cyclin dependent kinase like L	
AtCDKL.2	AT1G74330		
AtCDKL.3	AT1G18670		
AtCDKL.4	AT4G22940		
AtCDKL.5	AT5G44290		
AtCDKL.6	AT1G03740		
AtCDKL.7	AT5G50860		
AtCDKL.8	AT3G05050		
AtCDKL.9	AT1G54610		
AtCDKL.10	AT1G57700		
AtCDKL.11	AT1G09600		
AtCDKL.12	AT1G71530		
AtCDKL.13	AT4G10010		
AtCDKL.14	AT1G33770		
AtCDKL.15	AT1G53050		

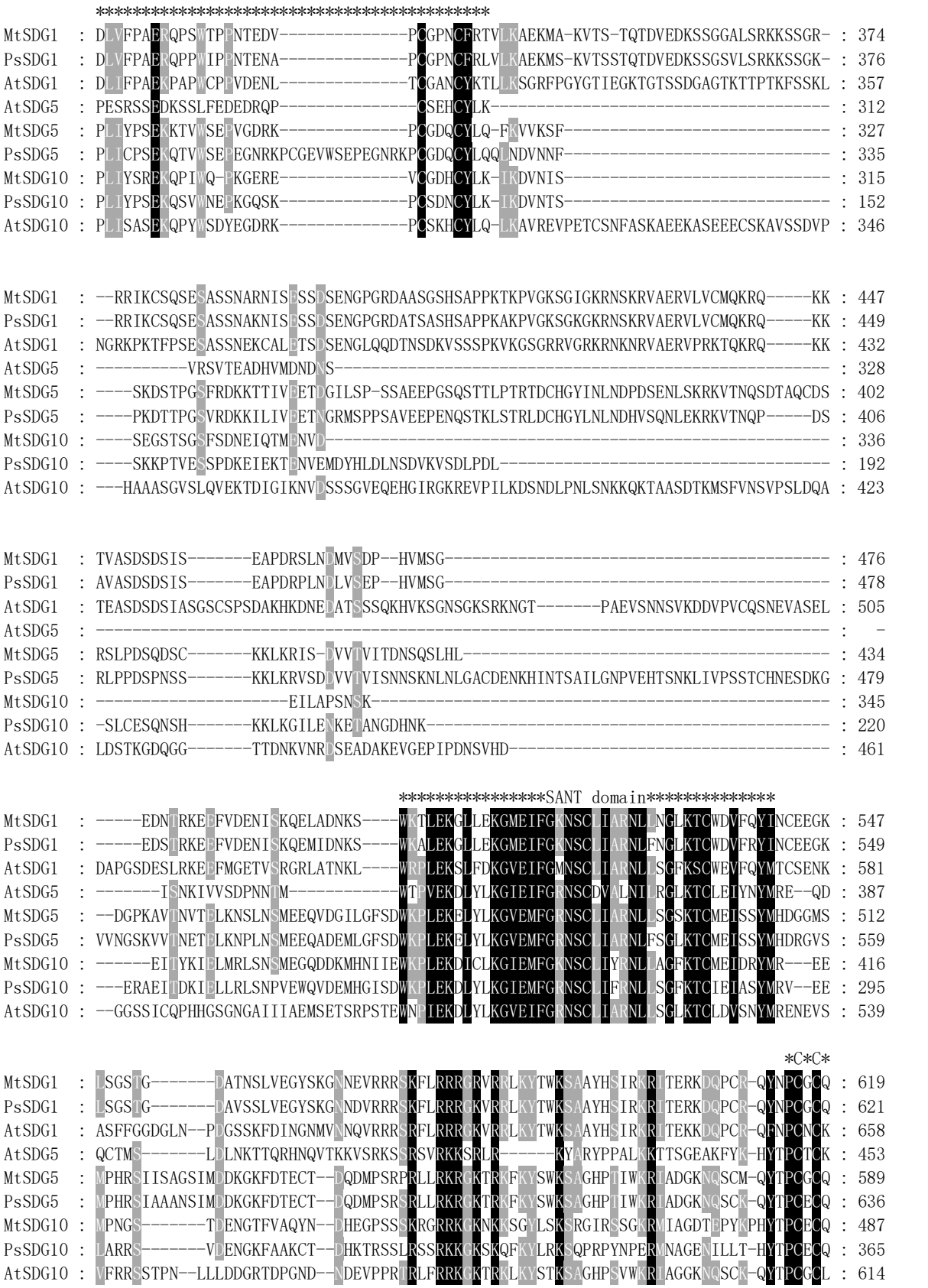
Table A2.4 Name and accession numbers of *A. thaliana* Cyclins.

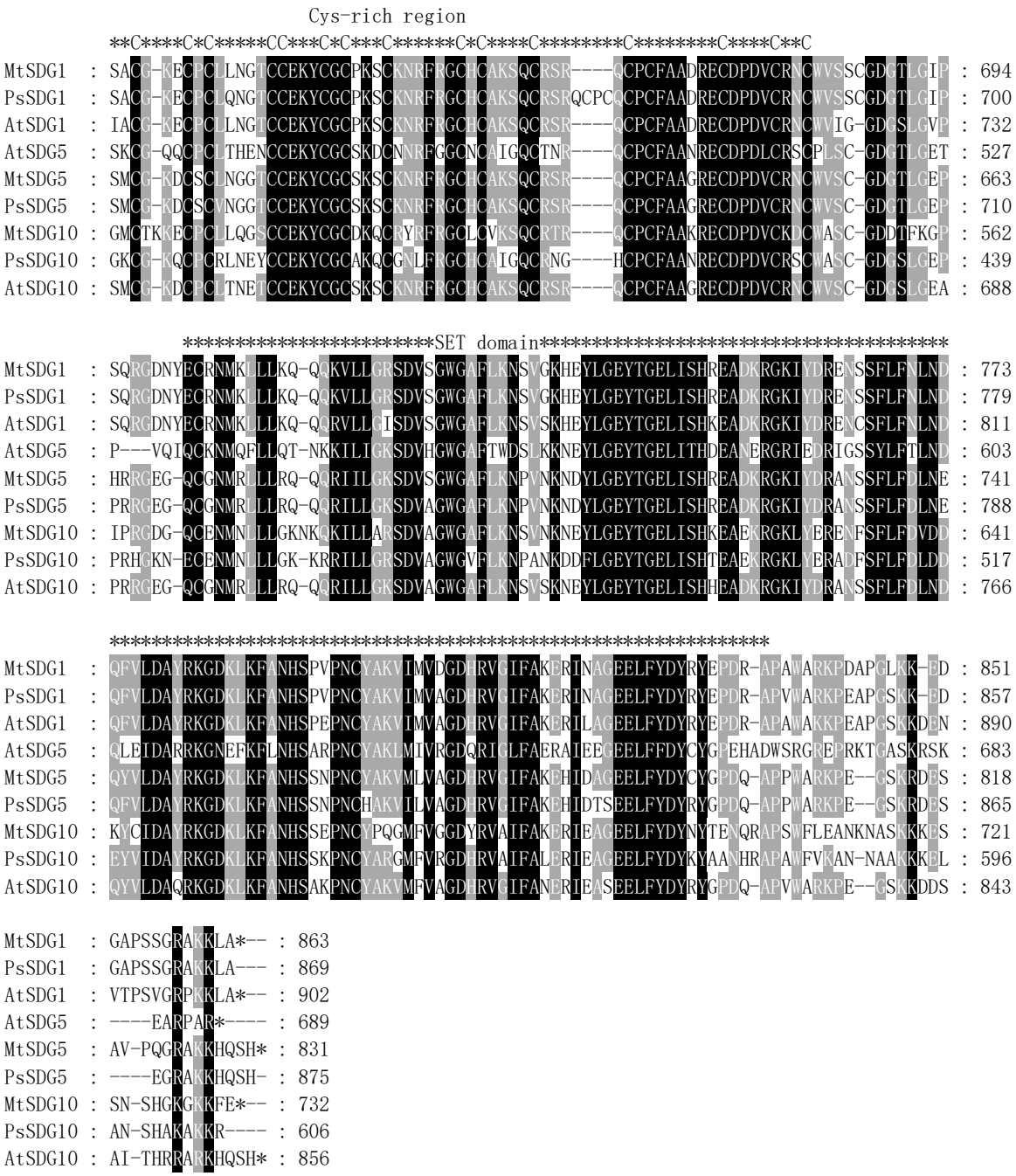
Name	Accession number	Groups	Source
AtCYCA1.1	AT1G44110	Cyclin A	Wang et al. 2004
AtCYCA1.2	AT1G77390		
AtCYCA2.1	AT5G25380		
AtCYCA2.2	AT5G11300		
AtCYCA2.3	AT1G15570		
AtCYCA2.4	AT1G80370		
AtCYCA3.1	AT5G43080		
AtCYCA3.2	AT1G47210		
AtCYCA3.3	AT1G47220		
AtCYCA3.4	AT1G47230		
AtCYCB1.1	AT4G37490	Cyclin B	
AtCYCB1.2	AT5G06150		
AtCYCB1.3	AT3G11520		
AtCYCB1.4	AT2G26760		
AtCYCB1.5	AT1G34460		
AtCYCB2.1	AT2G17620		
AtCYCB2.2	AT4G35620		
AtCYCB2.3	AT1G20610		
AtCYCB2.4	AT1G76310		
AtCYCB2.5	AT1G20590		
AtCYCB3.1	AT1G16330		
AtCYCC1.1	AT5G48640	Cyclin C	
AtCYCC1.2	AT5G48630		
AtCYCD1.1	AT1G70210	Cyclin D	
AtCYCD2.1	AT2G22490		
AtCYCD4.1	AT5G65420		
AtCYCD4.2	AT5G10440		
AtCYCD3.1	AT4G34160		
AtCYCD3.2	AT5G67260		
AtCYCD3.3	AT3G50070		
AtCYCD5.1	AT4G37630		
AtCYCD6.1	AT4G03270		
AtCYCD7.1	AT5G02110		
AtCYCH1.1	AT5G27620	Cyclin H	
AtCYCL1.1	AT2G26430	Cyclin L	
AtCYCT1.1	AT1G35440	Cyclin T	
AtCYCT1.2	AT4G19560		
AtCYCT1.3	AT1G27630		
AtCYCT1.4	AT4G19600		
AtCYCT1.5	AT5G45190		
SDS	AT1G14750	SDS	
AtCYCU1.1	AT3G21870	Cyclin U	
AtCYCU2.1	AT2G45080		
AtCYCU2.2	AT3G60550		
AtCYCU3.1	AT3G63120		
AtCYCU4.1	AT2G44740		
AtCYCU4.2	AT5G07450		
AtCYCU4.3	AT5G61650		
AtCYCJ18	AT2G01820	CyclinJ18	

Sequence A2.1. Alignment of the SDG proteins of group I of figure 5.1

Alignment of sequences for SDG proteins of group I in *A. thaliana* and their orthologues in *P. sativum* and *M. truncatula*. Alignments include full length predicted sequences for all the proteins. Shading represents degree of conservation, black = 100%, dark grey = 80%, light grey = 60%. Information on the gene names and accession numbers are presented in table 5.3.

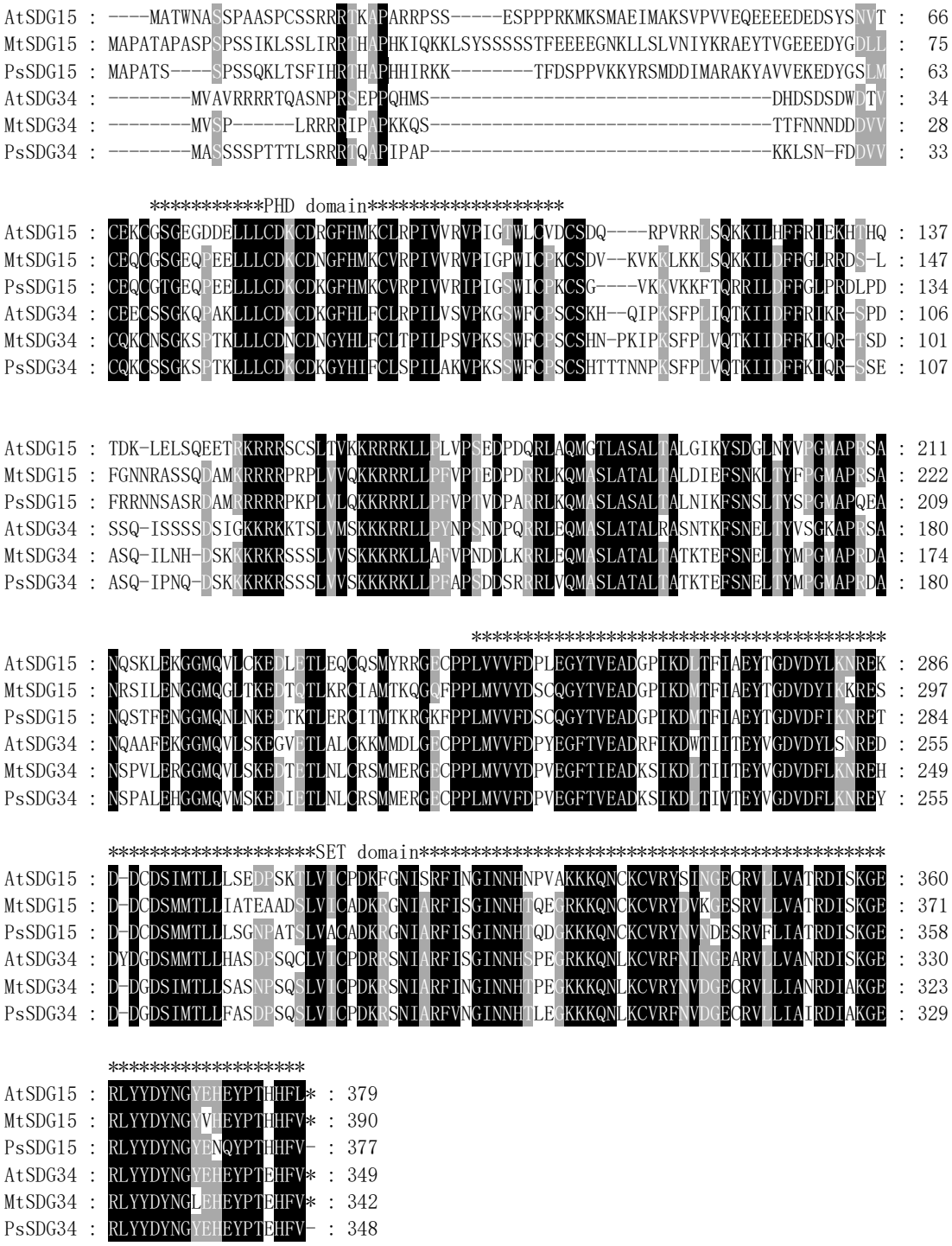
```
MtSDG1  : MASSSPPPPSPSSSSRSDPPLDPSTKKAEETSPAVKDVLAVIESLKKQVAAKRIVTVKTRVDENRQKLIATTNQLWKSSA :  80
PsSDG1  : MASKSPPPPSRSSSSRSEPPLDPSPTKMEDTGPAVKEVLAVIESLKKQVAAKRIVSVKNRVDENRQKLAGITNQLWKSAA :  80
AtSDG1  : MASEASPS---SSATRSEPPKDSP---AEERGPASKEVSEVIESLKKKLAADRCISIKKRIDENKKNLFAITQSFMRSSM :  74
AtSDG5  : MEK-------------------------ENHEDDGEGLPPEINQIKEQLEKERFLHIKRKFDLRYIPSVATHASHHQSFD :  55
MtSDG5  : MYSCSQVRLVASALNTNTNAEELV---EEPPNDDVGTLSNKINQLKKQLQAERIQSIKEKLQRVQKKLQCHTSGIMSAVS :  77
PsSDG5  : MVSKPND--------SPLILHQKLV---EEPADDAVGALSAKINQLKKQLQAERIEFIGEKLQRVQKKLQCHISGIMVAIS :  70
MtSDG10 : MASNIATS------ASRPRGQEQQ---GEATIEDPQILMNKLKKLKEKLQKERMETVQKKLQINEKNLQCELSKVMTTVS :  71
PsSDG10 : ------------------------------------------------------------------------------- :   -
AtSDG10 : MVTDDSNS------SGRIKSHVDDDDDGEEEEDRLEGLENRISELKRKLQGERVRSIKEKFEANRKKVDAHVSPFSSAAS :  74

                                                                                **********
MtSDG1  : ER-TCGIADTDRGLDLLSKRQKEAIDMHNGIR--AGNDDGESNGYNGDDHGSTAVLLGSNYAVKNAVRPIKLPEVKRLPP : 157
PsSDG1  : ERKTCDVADSDKSLDLLTKRQQEAIDMHNGVC--AGNDDGESNGNHGDDHGSTAVLLGSNVAVKNAVRPIKLPEVRRLPP : 158
AtSDG1  : ER-----GGSCKDGSDLLVKRQRDSPGMKSGID--ESNN-----NRYVEDGPASSGMVQGSSVPVKISLRPIKMPDIKRLSP : 144
AtSDG5  : LNQPAAEDDNGGDNKSLLSRMQNPLRHFSASSDYNSYEDQGYVLDEDQDYALEEDVPLFLDEDVPLLPSVKLPIVEKLPR : 135
MtSDG5  : TR-DSSQTEENTTRSILSSRMDRPLCKFYGFT--QGPGDRN-QGNQD----------------MSSATSIKIPRVERLPP : 137
PsSDG5  : TR-DSSQTEENKTCSILSSRMDHPLCKFDGFI--HVSGDKD-HSNQD----------------ISSVTSIKIPRIERLPP : 130
MtSDG10 : RY-DSS-IIGKDNVQTHSLKIEHPLEMYDRFP--RGLGNKYLHVVHD----------------VSFKKTFRLQRVEKIPH : 131
PsSDG10 : ------------------------------------------------------------------------------- :   -
AtSDG10 : SR---ATAEDNGNSNMLSSRMRMPLCKLNGFS--HGVGDRDYVPTKD----------------VISASVKLPIAERIPP : 132

          **********EZD1 domain*****************************************
MtSDG1  : YTTWIFLDRNQRMTEDQSVLGRRRIYYDQNGGEALICSDSEEELIDEEDEKR----EFVESEDFILRMTIREFGLSDVVL : 233
PsSDG1  : YTTWIFLDRNQRMTEDQSVLGRRRIYYDQNGGEALICSDSEEELIDEEDEKR----EFVEPEDFILRMTIREFGLSEAVL : 234
AtSDG1  : YTTWVFLDRNQRMTEDQSVVGRRRIYYDQTGGEALICSDSEEEAIDDEDEKR----DFLEPEDYIIRMTLEQLGLSDSVL : 220
AtSDG5  : SITWVFTKSSQLMAESDSVIGKRQIYYLNGEALELSSEEDQEDEEEDEDEIKKEKCEFSEDVDRFIWTVGQDYGLDDLVV : 215
MtSDG5  : YTSWIFLDRNQRMADDQSVVGRRRIYYDQRGSEALICSDSEEELTEPDGEKH----EFCDAEDRILCMAFEEHGLNEEVL : 213
PsSDG5  : YTSWIFLDRNQRMAEDQSVVGRRRIYYDQRGSEALICSDSEEELTEPEEVKH----EFCEAEDRILRMAFEEHGLNEEVL : 206
MtSDG10 : YTTWLHLIRNERMTKADAFSARRNIYYDQHAGETMICSDTDEEVQENKEVKR----DFSYGEDKLLWMAIEEYCLTDEVL : 207
PsSDG10 : ---------------------MINIVVKQ-------------EECKENREAKH----DFSQGEDQILWMAFEEHDFTEEAL :  43
AtSDG10 : YTTWIFLDRNQRMAEDQSVVGRRQIYYEQHGGETLICSDSEEEP-EPEEEKR----EFSEGEDSIIWLIGQEYGMGEEVQ : 207

                                                **********EZD2 domain*******
MtSDG1  : E-ILAQCFSRKTSDIKVRYETFCNEDNSGEDSKNGDAQDNSQIDDSFLEKDLEAALDSFDN-LFCRRCRVFDCRLHGCSQ : 311
PsSDG1  : E-SLAQCFSRKTSDIKARYETFCNEDNAGGDSKNGDAQDTSQSDNSFLEKDLEAALDSFDN-LFCRRCRVFDCRLHGCSQ : 312
AtSDG1  : A-ELASFLSRSTSEIKARHGVLMKEKEVSESGDN--------QAESSLLNKDMEGALDSFDN-LFCRRCLVFDCRLHGCSQ : 291
AtSDG5  : RRALAKYLEVDVSDILERYNELKLKN-----------DGTAGEASDLTSKTITTAFQDFADRRHCRRCMIFDCHMHEKYE : 284
MtSDG5  : N-VVSKYVGGTSLEIQERYKSIRGSNIGRLDQHPKSSGEHEFPMSMYLEKNLSDALDSLDN-LFCRRCLIFDCRLHGCSQ : 291
PsSDG5  : D-VVCKHVGGTSLEIQERYKSIRENTTGRLGQHSKSLGEHDSPMSMYLEKTLSDALDSLDN-LFCRRCLIFDCRLHGCSQ : 284
MtSDG10 : S-IVQSYIGGTTAEIEERYKYLKEK-----SMLSKDSRENASNSGLCLDKSLSEALSTFDH-FFCRRCLIFDCPLHGCSQ : 280
PsSDG10 : S-VVQRCIGGTCSEIQERYKYLKEK-----DLHAKNSRESESITGICLDKSLNASLSTFDH-LFCRRCMIYDCPLHGCSQ : 116
AtSDG10 : D-ALCQLLSVDASDILERYNELKLKD----KQNTEEFSNSGFKLGISLEKGLGAALDSFDN-LFCRRCLVFDCRLHGCSQ : 281
```


MtSDG1
PsSDG1
AtSDG1
AtSDG5
MtSDG5
PsSDG5
MtSDG10
PsSDG10
AtSDG10
SANT domain
*C*C*


Cys-rich region
C**C*C*****CC***C*C***C******C*C****C********C********C****C**C
MtSDG1 : SACG-KECPCLLNGTCCEKYCGCPKSCKNRFRGCHCAKSQCRSR----QCPCFAADRECDPDVCRNCWVSSCGDGTLGIP : 694
PsSDG1 : SACG-KECPCLQNGTCCEKYCGCPKSCKNRFRGCHCAKSQCRSRQCPCQCPCFAADRECDPDVCRNCWVSSCGDGTLGIP : 700
AtSDG1 : IACG-KECPCLLNGTCCEKYCGCPKSCKNRFRGCHCAKSQCRSR----QCPCFAADRECDPDVCRNCWVIG-GDGSLGVP : 732
AtSDG5 : SKCG-QQCPCLTHENCCEKYCGCSKDCNNRFGGCNCAIGQCTNR----QCPCFAANRECDPDLCRSCPLSC-GDGTLGET : 527
MtSDG5 : SMCG-KDCSCLNGGTCCEKYCGCSKSCKNRFRGCHCAKSQCRSR----QCPCFAAGRECDPDVCRNCWVSC-GDGTLGEP : 663
PsSDG5 : SMCG-KDCSCVNGGTCCEKYCGCSKSCKNRFRGCHCAKSQCRSR----QCPCFAAGRECDPDVCRNCWVSC-GDGTLGEP : 710
MtSDG10 : GMCTKKECPCLLQGSCCEKYCGCDKQCRYRFRGCLCVKSQCRTR----QCPCFAAKRECDPDVCKDCWASC-GDDTFKGP : 562
PsSDG10 : GKCG-KQCPCRLNEYCCEKYCGCAKQCGNLFRGCHCAIGQCRNG----HCPCFAANRECDPDVCRSCWASC-GDGSLGEP : 439
AtSDG10 : SMCG-KDCPCLTNETCCEKYCGCSKSCKNRFRGCHCAKSQCRSR----QCPCFAAGRECDPDVCRNCWVSC-GDGSLGEA : 688
**********SET domain**********
MtSDG1 : SQRGDNYECRNMKLLLKQ-QQKVLLGRSDVSGWGAFLKNSVGKHEYLGEYTGELISHREADKRGKIYDRENSSFLFNLND : 773
PsSDG1 : SQRGDNYECRNMKLLLKQ-QQKVLLGRSDVSGWGAFLKNSVGKHEYLGEYTGELISHREADKRGKIYDRENSSFLFNLND : 779
AtSDG1 : SQRGDNYECRNMKLLLKQ-QQRVLLGISDVSGWGAFLKNSVSKHEYLGEYTGELISHKEADKRGKIYDRENCSFLFNLND : 811
AtSDG5 : P---VQIQCKNMQFLLQT-NKKILIGKSDVHGWGAFTWDSLKKNEYLGEYTGELITHDEANERGRIEDRIGSSYLFTLND : 603
MtSDG5 : HRRGEG-QCGNMRLLLRQ-QQRIILGKSDVSGWGAFLKNPVNKNDYLGEYTGELISHREADKRGKIYDRANSSFLFDLNE : 741
PsSDG5 : PRRGEG-QCGNMRLLLRQ-QQRILLGKSDVAGWGAFLKNPVNKNDYLGEYTGELISHREADKRGKIYDRANSSFLFDLNE : 788
MtSDG10 : IPRGDG-QCENMNLLLGKNKQKILLARSDVAGWGAFLKNSVNKNEYLGEYTGELISHKEAEARGKLYERENFSFLFDVDD : 641
PsSDG10 : PRHGKN-ECENMNLLLGK-KRRILLGRSDVAGWGVFLKNPANKDDFLGEYTGELISHTEAEARGKLYERADFSFLFDLDD : 517
AtSDG10 : PRRGEG-QCGNMRLLLRQ-QQRILLGKSDVAGWGAFLKNSVSKNEYLGEYTGELISHHEADKRGKIYDRANSSFLFDLND : 766
****
MtSDG1 : QFVLDAMRKGDKLKFANHSPVPNCYAKVIMVDGDHRVGIFAKERINAGEELFYDYRYEPDR-APAWARKPDAPGLAK-ED : 851
PsSDG1 : QFVLDAMRKGDKLKFANHSPVPNCYAKVIMVAGDHRVGIFAKERINAGEELFYDYRYEPDR-APVWARKPEAPGSAK-ED : 857
AtSDG1 : QFVLDAMRKGDKLKFANHSPEPNCYAKVIMVAGDHRVGIFAKERILAGEELFYDYRYEPDR-APAWAKKPEAPGSKKDEN : 890
AtSDG5 : QLEIDARRKGNEFKFLNHSARPNCYAKLMIVRGDQRIGLFAERAIEEGEELFFDYCYGPEHADWSRGREERKTGASKRSK : 683
MtSDG5 : QYVLDAMRKGDKLKFANHSSNPNCYAKVMLVAGDHRVGIFAKEHIDAGEELFYDYCYGPDQ-APPWARKPE--GSKRDES : 818
PsSDG5 : QFVLDAMRKGDKLKFANHSSNPNCHAKVILVAGDHRVGIFAKEHIDTSEELFYDYRYGPDQ-APPWARKPE--GSKRDES : 865
MtSDG10 : KYCIDAMRKGDKLKFANHSSEPNCYPQGMFVGGDYRVAIFAKERIEAGEELFYDYNYTENQRAPSWFLEANKNASKKKES : 721
PsSDG10 : EYVIDAMRKGDKLKFANHSSKPNCYARGMFVRGDHRVAIFALERIEAGEELFYDYKYAANHRAPAWFVKAN-NAAKKKEL : 596
AtSDG10 : QYVLDAQRKGDKLKFANHSAKPNCYAKVMFVAGDHRVGIFANERIEASEELFYDYRYGPDQ-APVWARKPE--GSKKDDS : 843
MtSDG1 : GAPSSGRAKKLA*-- : 863
PsSDG1 : GAPSSGRAKKLA--- : 869
AtSDG1 : VTPSVGRPKKLA*-- : 902
AtSDG5 : ----EARPAR*---- : 689
MtSDG5 : AV-PQGRAKKHQSH* : 831
PsSDG5 : ----EGRAKKHQSH- : 875
MtSDG10 : SN-SHGKGKKFE*-- : 732
PsSDG10 : AN-SHAKAKKR---- : 606
AtSDG10 : AI-THRRARKHQSH* : 856

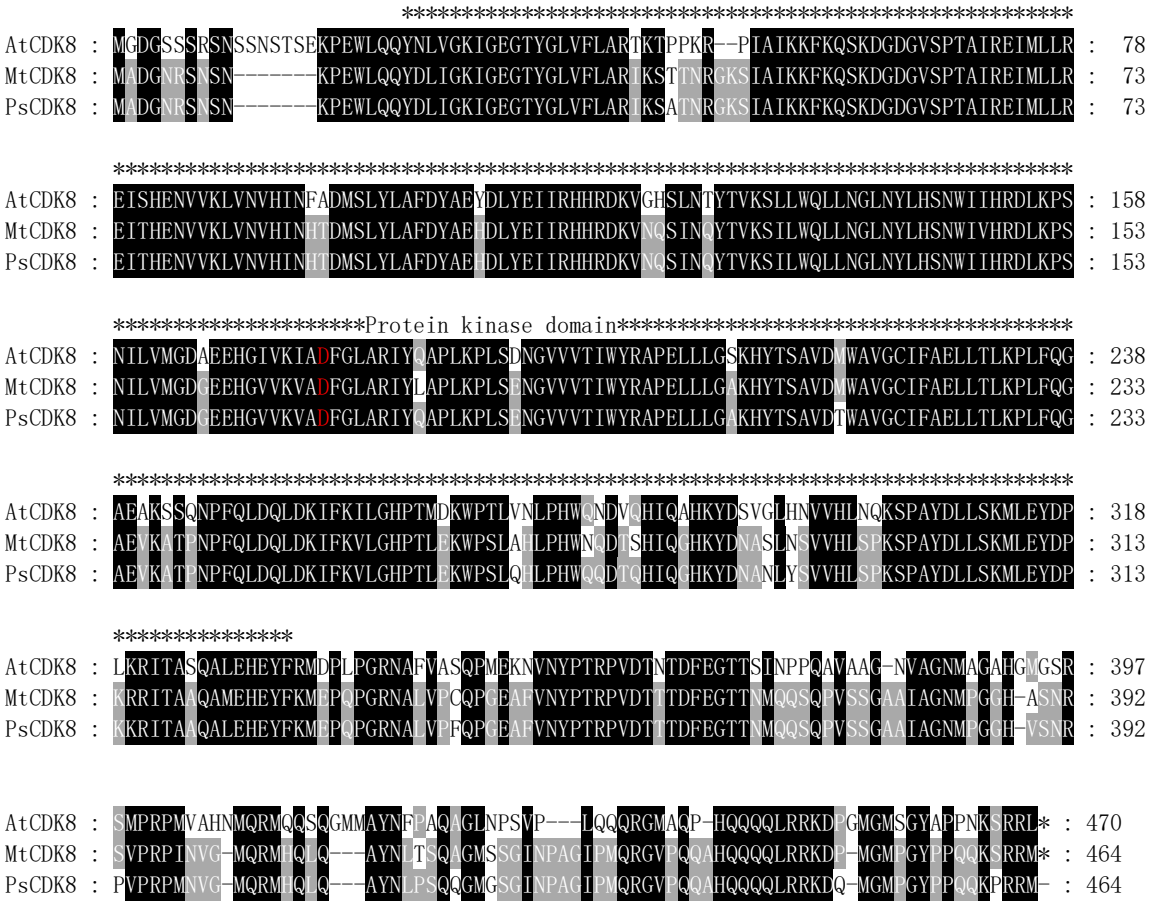
Sequence A2.2. Alignment of the SDG protein of group IV of figure 5.1

Alignment of sequences for SDG proteins of group IV having SET domain in *A. thaliana* and their orthologues in *P. sativum* and *M. truncatula*. Alignments include full length predicted sequences for all the proteins. Shading represents degree of conservation, black = 100%, dark grey = 80%, light grey = 60%. Information on the gene names and accession numbers are presented in table 5.3.



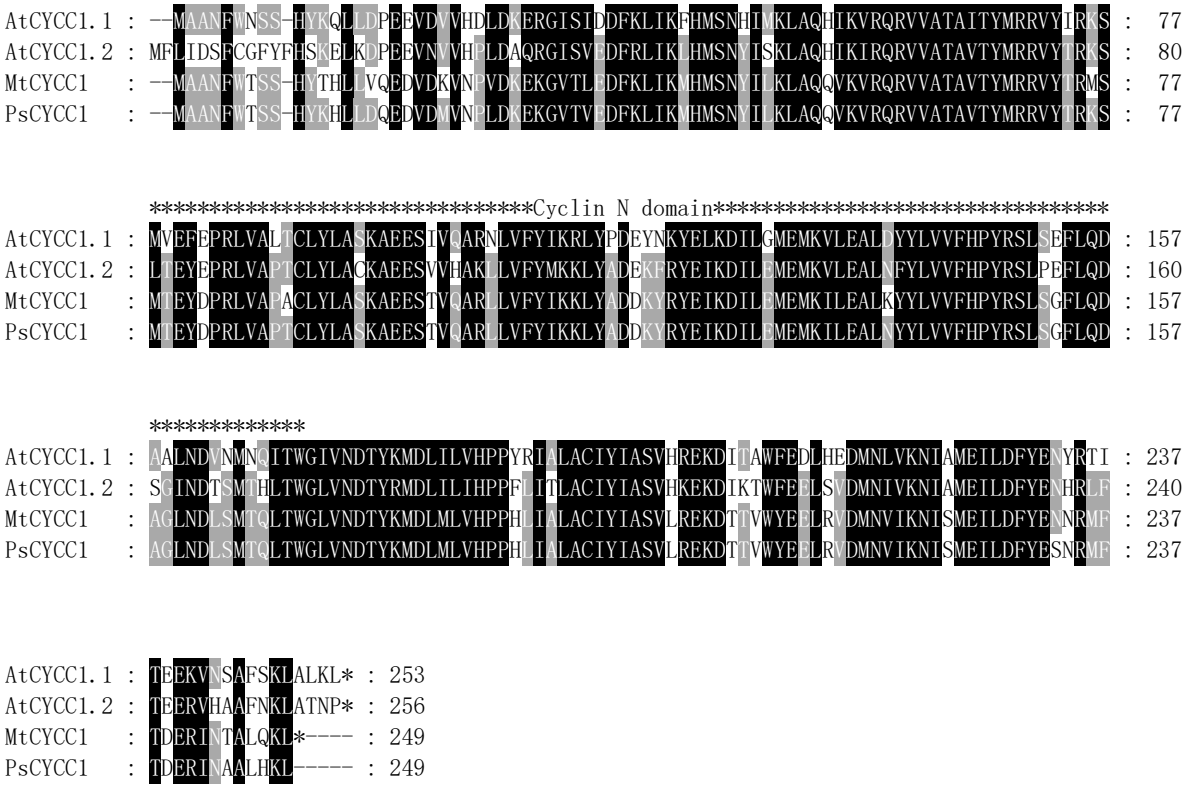
Sequence A2.3. Alignment of Cyclin dependent kinase 8 proteins for Figure 5.5

Alignment of sequences for Cyclin dependent kinase 8 (CDK8) proteins from *A. thaliana* and its orthologues in *M. truncatula* and pea. The conserved aspartic acid marked in red color is crucial for the kinase activity of CDK8. Alignments include full length predicted sequences for all the proteins. Shading represents degree of conservation, black = 100%, dark grey = 80%, light grey = 60%. Information on the gene accession numbers are presented in Figure 5.5.



Sequence A2.4. Alignment for Cyclin C proteins from Figure 5.7

Alignment of sequences for Cyclin C (CYCC1) proteins from *A. thaliana* and their orthologues in *M. truncatula* and pea. Alignments include full length predicted sequences for all the proteins. Shading represents degree of conservation, black = 100%, dark grey = 80%, light grey = 60%. Information on the gene accession numbers are presented in Figure 5.7.



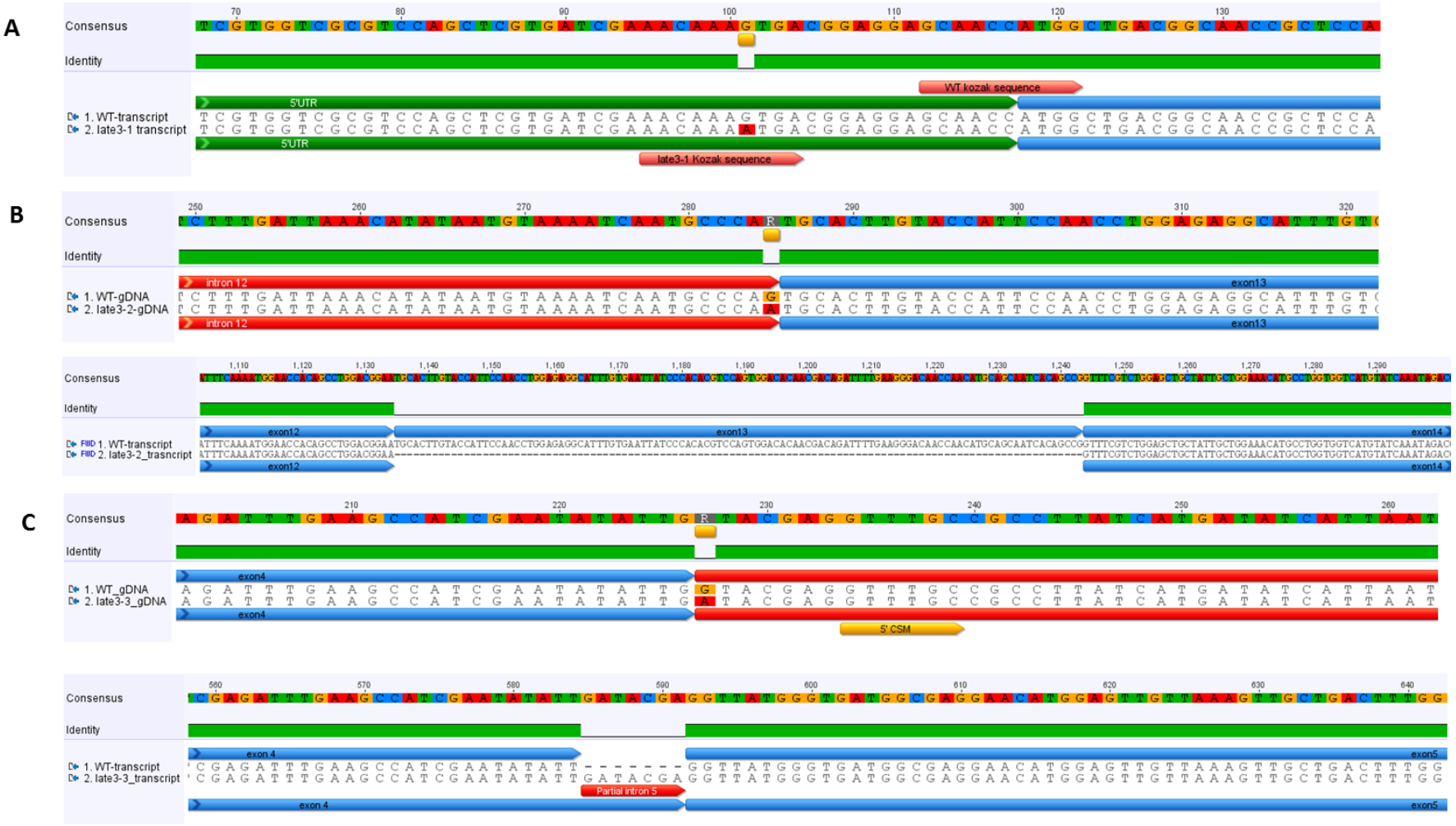


Figure A2.1. Sites and consequences of mutation obtained by direct sequencing of *PsCDK8* gene of various mutant alleles of *late3* mutant. **(A)** Alignment of WT and *late3-1* transcript. **(B)** Alignment of WT and *late3-2* genomic DNA (gDNA) and transcript. **(C)** Alignment of WT and *late3-3* gDNA and transcript. CSM= cryptic splice motif.

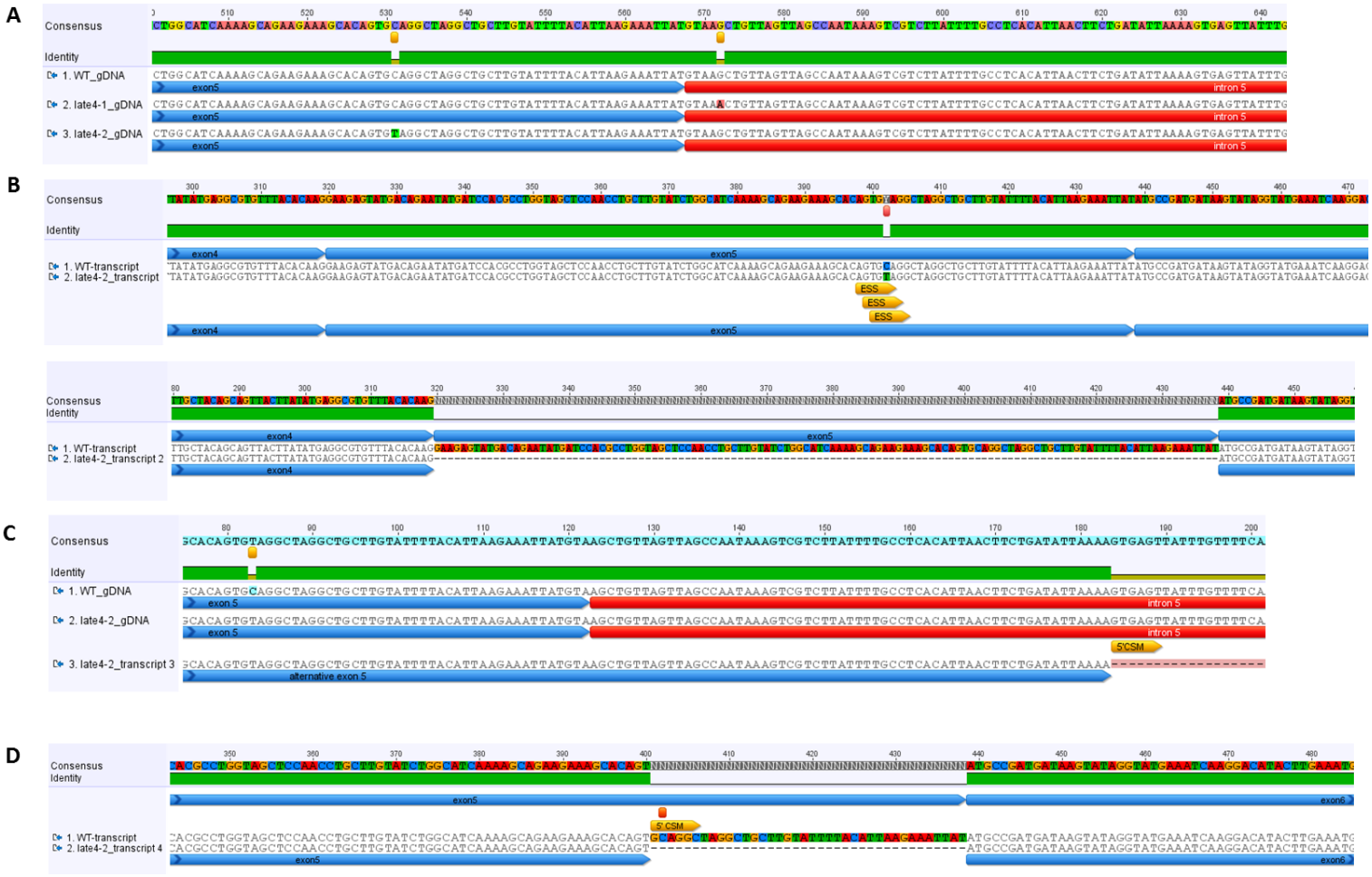


Figure A2.2. Sites and consequences of mutation obtained by direct sequencing of *PsCYCC1* gene of various alleles of *late4* mutant. **(A)** Alignment of WT, *late4-1* and *late4-2* gDNA. **(B)** Alignment of WT and *late4-2* transcript showing presence of putative exon splicing silencer motifs (ESS) in exon5 of *late4-2* genotype introduced due to the mutation which may have resulted in generation of *late4-2* transcript type 2 by skipping entire exon 5. **(C)** Alignment of WT and *late4-2* gDNA and *late4-2* transcript type 3 showing partial retention of intron 5. **(D)** Alignment of WT and *late4-2* transcript type 4 showing partial deletion of exon 5. **(continued next page)**

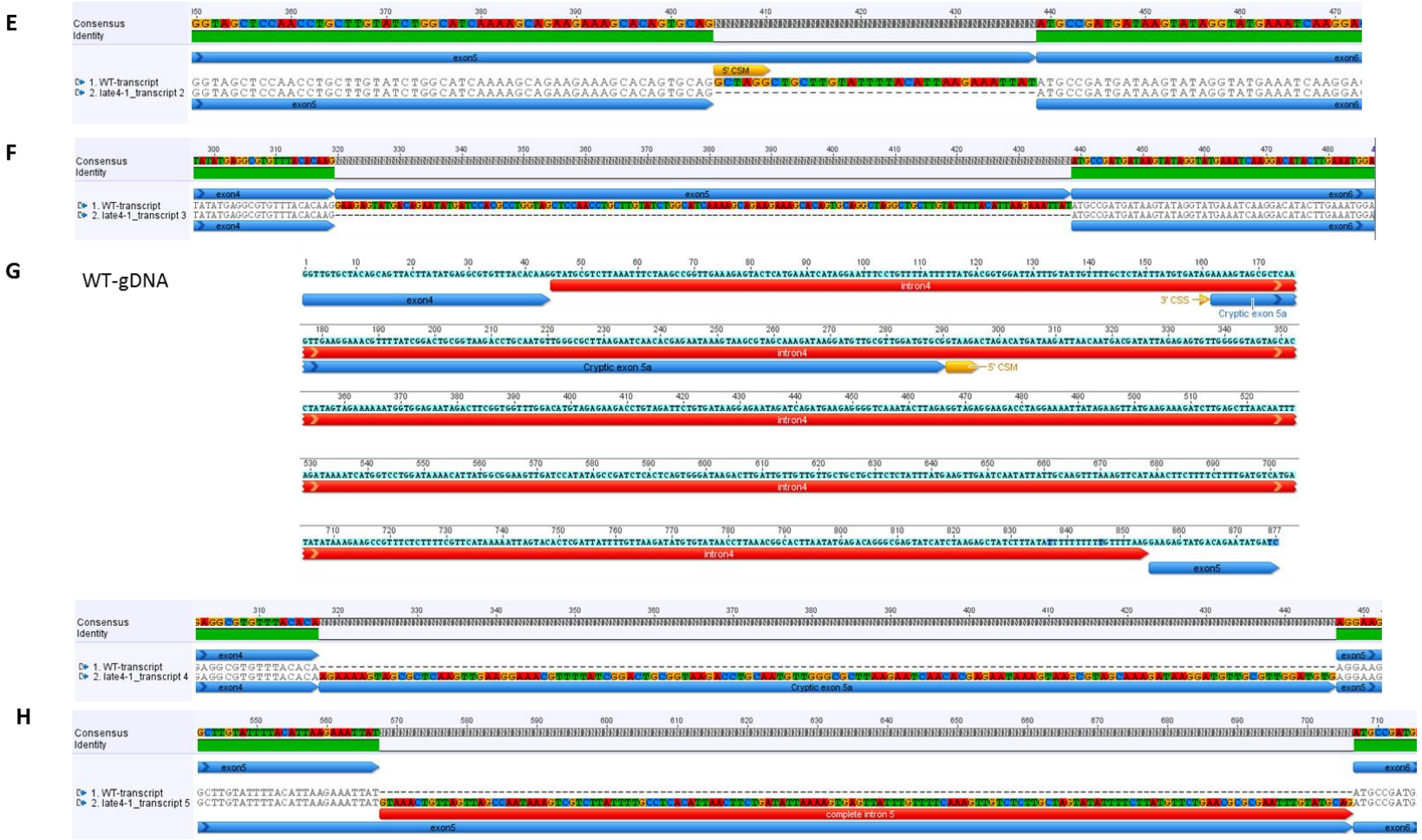


Figure A2.2. (continued) (E) Alignment of WT and *late4-1* transcript type 2 showing partial deletion of exon 5. **(F)** Alignment of WT and *late4-1* transcript type 3 showing skipping of exon 5. **(G)** Genomic DNA (gDNA) of WT around intron 4; alignment of WT and *late4-1* transcript type 4 showing activation of cryptic exon 5a and also similar situation shown in figure A2.2 F. **(H)** Alignment of WT and *late4-1* transcript type 5 showing complete intron 5 retention and also similar situation shown in figure A2.2 G. The ESS shown in E are predicted from the FAS-ESS web server (http://genes.mit.edu/fas-ess/). CSM= cryptic splice motif, CSS= cryptic splice site, ESS = Exon splicing silence motif.

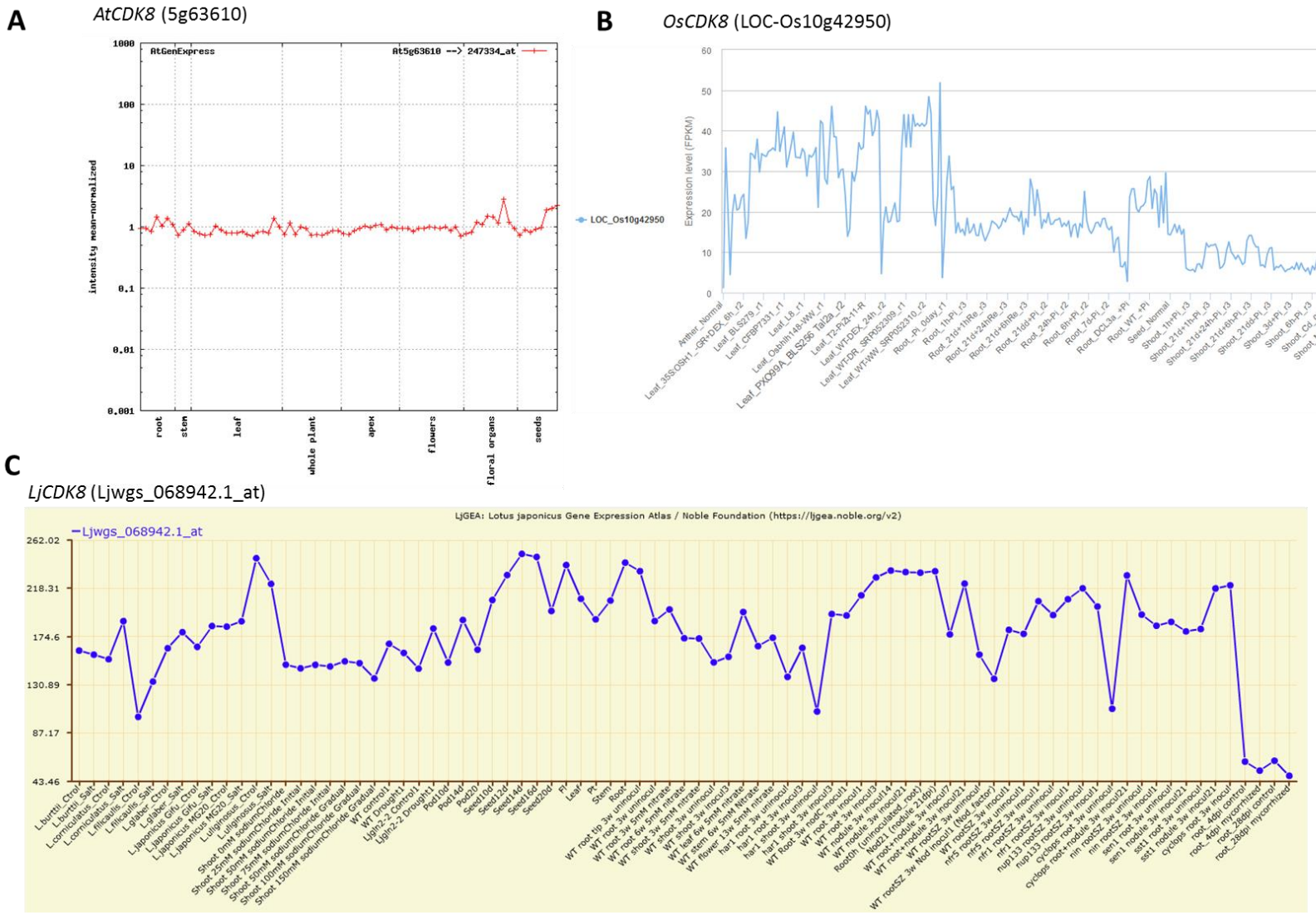
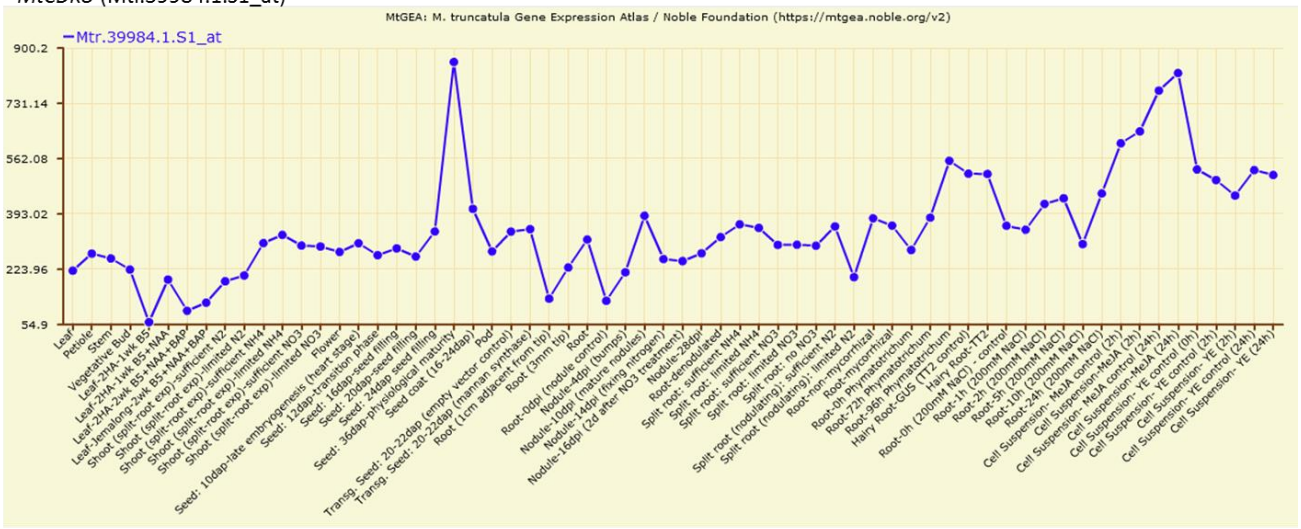


Figure A2.3. Expression atlas for *CDK8* gene in *A. thaliana, O. sativa, L. japonicus*, *M. truncatula* and *P. sativum* **(continued next page)**

D *MtCDK8* (Mtr.39984.1.S1_at)



E *PsCDK8* (PsCam048317)

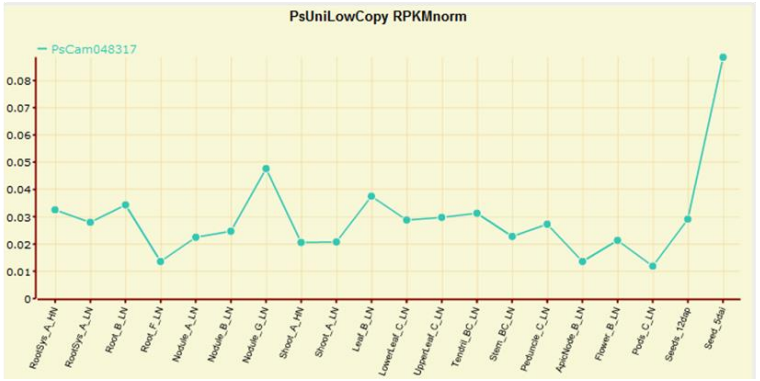


Figure A2.3. (continued) (A) *A. thaliana* data sourced from AtGenExpress Visualization Tool (http://jsp.weigelworld.org/expviz/expviz.jsp). **(B)** *O. sativa* from Rice Expression Database - IC4R (http://expression.ic4r.org/) **(C)** *L. japonicus* from *Lotus* Gene Expression Atlas (https://ljgea.noble.org/v2/). **(D)** *M. truncatula* from *Medicago* Gene Expression Atlas (https://mtgea.noble.org/v2/) and **(E)** *P. sativum* from Pea RNA seq gene atlas (http://bios.dijon.inra.fr/FATAL/cgi/PsUniLowCopy.cgi)

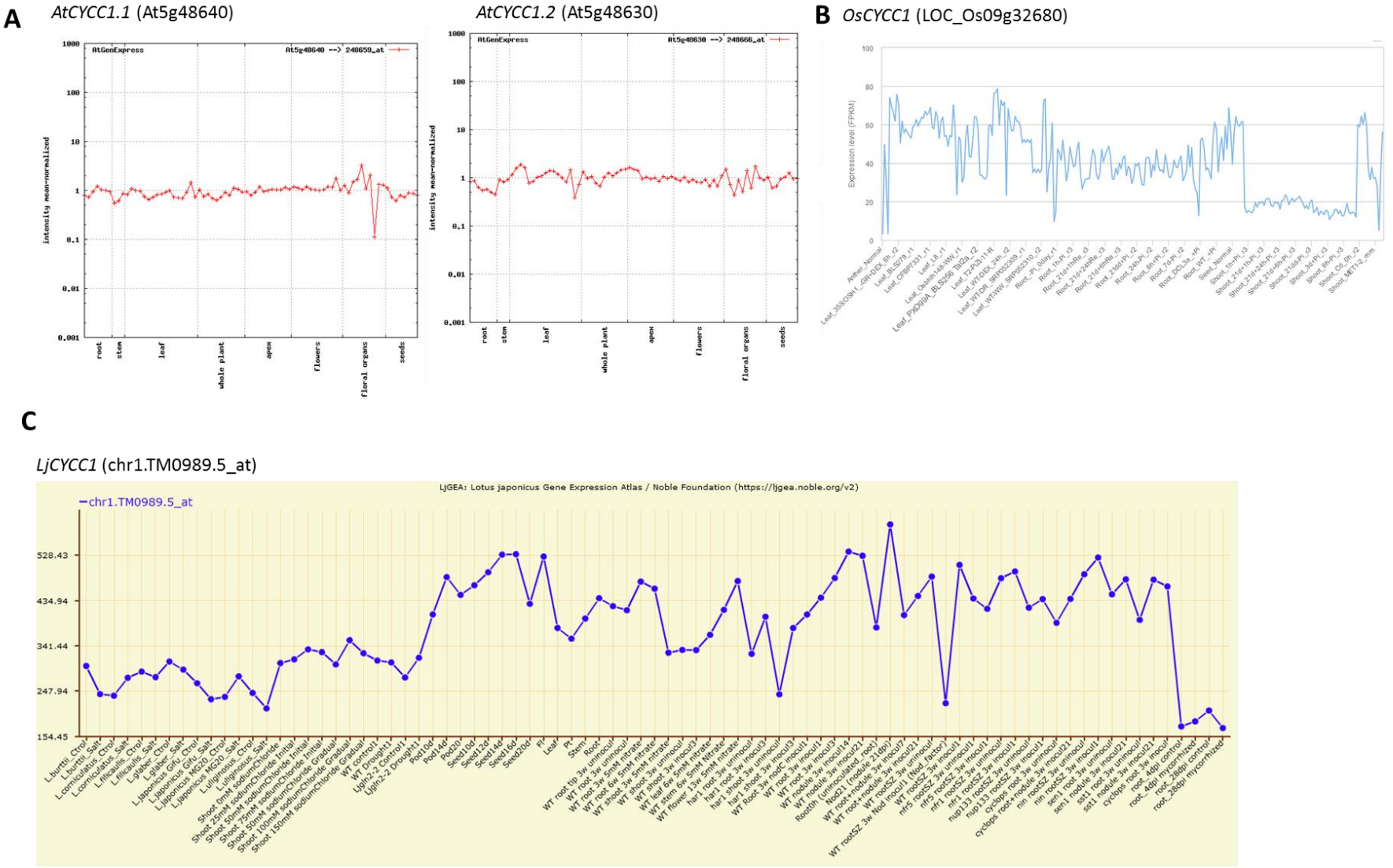
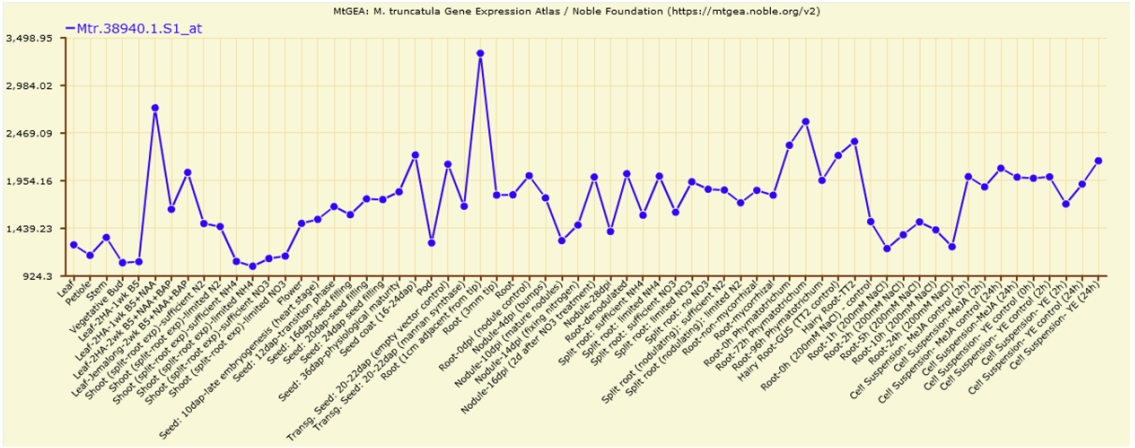


Figure A2.4. Expression atlas for *CYCC1* gene in *A. thaliana, O. sativa, L. japonicus*, *M. truncatula* and *P. sativum* **(continued next page)**

D

MtCYCC1 (Mtr.38940.1.S1_at)



E

PsCYCC1 (PsCam050605)

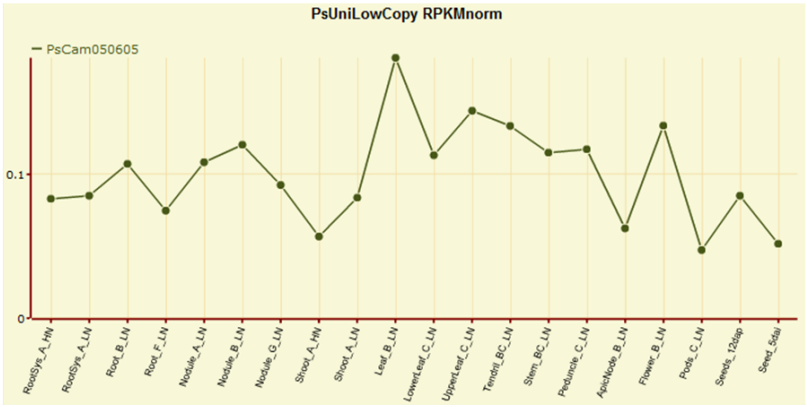


Figure A2.4. (continued) (A) *A. thaliana* data sourced from AtGenExpress Visualization Tool (http://jsp.weigelworld.org/expviz/expviz.jsp). **(B)** *O. sativa* from Rice Expression Database - IC4R (http://expression.ic4r.org/) **(C)** *L. japonicus* from *Lotus* Gene Expression Atlas (https://ljgea.noble.org/v2/). **(D)** *M. truncatula* from *Medicago* Gene Expression Atlas (https://mtgea.noble.org/v2/) and **(E)** *P. sativum* from Pea RNA seq gene atlas (http://bios.dijon.inra.fr/FATAL/cgi/PsUniLowCopy.cgi)

Appendix 3- Sequences for chapter 6

Appendix sequence A3.1. *MtCDK8* (Medtr3g096960), -2000 to +1 bp of Translation initiation site as 5'UTR is not given for this gene in *M. truncatula* Mt4.0v1 database for determination of promoter sequence. The red colored letters have been inferred as 5'UTR (117 bp) of *MtCDK8* through RNA sequencing analysis of *PsCDK8* (previous chapter). The entire sequence was used for getting predicted TF binding shown in table 6.4.

1 bp

TTTCCCCACTATCTTTCATCACCATTCTCTCTCTCTCCTCCATGCTCCAATCCACCACTAGCAATGTTCATGGTCTCTGAATAT
TTGTTTCCCTTTCTTGAATATTCCAACAAAAATTCCAGTTACAAATTTGTGTGTTTTTCAATTTTCCATCCAATTGAATTTCAA
ACCAACACCAAGAGTTTCTCCATCAAGTAGATTGTCTACAATGTTCATGACTTCATGATCTCTCTCTCTTCTAGTTCATAACT
TCATATTCTGTCCTCTTTTAGATTGTCTACAATTATTTTTTATCTCTAGTCTGGTTCATCACAAGAAAAACTATCATTAATTGA
TTTTCTAAACATATTTCTTTTTCATTTGTTTGATTTCGTAGCTGGTTCTTATCCAATACGAAAAAATAGAACATATAATAAGT
GATTTTAGATACTTCTTTCAAAAAAAGTGATTTTAGATACTAAAATGTTCAATTGGATTCTACCAAAAAAACCAGAATTCTA
CCAAAAAAATAATGTTAATTAAAAGAAAAAAAAATTCCAGATCTTAACAGAAATGTAAGACAGTAAGAGTGTAAAAGAG
AAAGGAAAAAAATATGGTGTAGGTAGTCCATGATAAAATTAATACTCTTCTTTAATCTAATGGCTGAAAAAGACTTTTCCA
TGGTGGGAAAGAATTCTTCCTACCCACCCTAAATGCCACCTTAACTGATAGTTGACTAACAAAAGTTAATATAGCAGTGTT
CAATTGCGACCATCTGAATGCCGACATAACATGTCATGTCATACTTTTTATTTCCATCGTTTATTATTAATTATGAATACTAT
TATTAGTAAGAGGTAAATCAACAATGGAAATTGTACCAAAAAAAAAAAAAAATCAACAATGGAAATATTATTGAGATGATG
ATGTGAGCCATAGTCTATCTCCCTTCTTCTTCTCAGTATGAATTTGAGAAATTTGGAGTTGTTGACAAGATATTTTGATATTT
TTTTGTTTGTCTGAAACAATGAGTGATATTTTCGTTTTCAAAATCAACTTTAATTTAAGTTAGCAATATTCTATCCTATTCACC
ATGGACAAGTCAGCAAAAACGCATCACTTCATCAAAGAAAAAGGGACCATAATCAAATTAAAAATATTAGTAATTACAAT
GTTTCTTTATAACAATTAATAATTAAAAGGACTAACCATCAGGTGACACATTTCTTAAAAAAATAATAATCAAGTGACACAT
ATTTACGAAGACGTTTTAATTTTAAACCAAGATAAATCATTGATAAAAGTTGTCCTTCATATAAAACTATCAAAAATTAGTT
TTAATATTAGTTAATTATTTTAGTTTGTGATTGAATTAAGACCGAATTGAGTATATTATAAAAGTAAAAATAAATTTGTTAA
AAAAATAAAATAAAAATAAAAACACTTTACTAAAGTAAAAATAAAAGAGTAAATAGACAATTACCCATAAAATTGTAAGTT
TTTTCAATTATCCTCTGAAATTAACAAAACTTCAATTAAATTTCACAACGTTAATTAATTTACGTCAAATTTTTCTGTTAGTAA
ACATGACGTTTTGCAAATACCCCTAAAGTTTTGCACTTATGTGCAAAATGTCCCCTGAACTTAAAAATTTATATTATTTTTTT
TATAAGTGAAACACTAGAATAAAGAAAAATACACACCATAAGCACAAACAAACTTTAGGGGATATTTCGCAGATAAGTGT
CAAATTTTAGAGGGTATTTACAAAACGTCATTTACTAACAGAAAAATTTGAGGGAGAGAGATAAAATGATTAACGTTGTA
AAATTTTGGAGGTAATTGAAGTTTTGTTAATGGTAAGAGGGTAATTGATAAGAGGGTAATTGATGAAACTTAATAAAAAA
ACACTTTGCTTATGCAGGAGCGGACAAATATCCAAACGGTACATTCAAACAGAACCTTAACCAGTTTCCGCTTGGTGATTG
TGTCGAGTTCGTG**A**

2000 bp

Appendix sequence A3.2. *MtCYCC1* (Medtr7g055650), -2000 to +1 bp of Translation initiation site as 5'UTR is not given for this gene in *M. truncatula* Mt4.0v1 database for determination of promoter sequence. The red colored letters have been inferred as 5'UTR (117 bp) of *MtCYCC1* according to *PsCYCC1* transcript sequence given in pea transcriptome database (previous chapter). The entire sequence was used for getting predicted TF binding shown in table 6.5.

1 bp

TAACGTTATGCAATTTCGGGGGGTAATTGAAGTTTTGTTAATTTCAGGGGGTAATTGACAAAACCTACAATTTCAGGGGG

AAATTGACTATTTACTCATTTAAGATTGGATTGGAATTCAGTATCTCCGATTAAACTTAGATAAATCAGAGTTAATTTATTT

AGTTATCGCTAACACACAAGAGAATTGCTGTTCACACATTTGGTTCGGCTGAGAAACACGAGTAAATTACTCAAATGCCCC

TCGACAGTTAACTGCCAAAATTTTAAAGAATCGGCTACAGTTAACTGCCGAGGAAAACTTTGGCAGTTAACTGCCGAGGA

TTTTCAAACATGAAATTAAATAAGGAAATGCATATAAATTTTCCTTAATTAAATTTAATTCTTTAATCCTTTTAAGGAATTAA

ACATAATTATTTTTATGATTTCATTCATATGATTTTTGTTGCCATAAATGAAGAAATATGAGAGATGAATAATATCTGCAAT

CGGTAGTTAATGGACCTAAAATGATGATTCAATAAAAAAATGCAACAACAAAAACGTTACAAATGATAGAATAATTATATT

TGTAATTATTACTCCCTCCGTCCCTAAATAAATGACCTAGTTGATTCTGACACACATACCAATGCATATGTTTTATCTTTGAT

ATCTTCAATTCTCTACTAAAAAAAATTATAAAAATTTAATATTTTAAAAATACTCATCGAGACGAATCCAGCAACATCTTAC

ATGATATTATTTATCTTTGTGAATTAGTATAAAAGTATGGTCAAAGTATGTCAAATCAATAATGTATATTGTCAACTGGGTC

ATTTATTAAGGGACGGAGGGAGTATTAATTAATTTCATTGATGTCTGTCATTATTCATAGGTTTGACTGACCTATTATTACT

TTATACTAATATTTCAATATTATTACTTTAATTAACATGAAAAACGCAACAACAAAAAATAAAAATGTTTGCCATCGGTAAT

TAATGGACCTAAAATGGTGATTCAATAAAAAAAAATGCAACAACAAAAAAACATTATAAATGATAGAATAATTATATTTAT

AATTAATTTCATTGATGTGTGTCATTATTAATAGGTTTGACTAACCTATTATTACTTTATAATATTTTAGATATTATCATCTGT

CTTTCAAATAATTATATTTGTAATTATTCTATTAATAAAACCTTTTTTTTATAGAAATAAAACCTCTTTTCTATTACTAGTAAG

AAACGTGATCGAATACTACAATTGTTTGAAACGTACAAAATAATGATTGAAACATTCAAAATCAATAGATATTATTAATGA

TAATTTATGGTTTTTGTCCAAAAAAAAAAGGTTTGAGGTTTCCTCGTCAGTTAACTGCCGAAGTTTTCAGTTTTCGTCGGTA

ATTTACTGCCGTAGTTTTTCTCGGAAGTTACTGCAGCCGGTTTCATAATTCGACAGTTAACTGCCGAGGGATATTTCGGTA

AATTACTAGTGTGACACAATCAATGCATATGTGTCAACCGCAATTCTCAACACATAAAACAAAGAAGCACGATAATCAATG

ATCCTAGATTTAAGAACAGAAACAAAAGACACGAACTACTAGTTTCCCCTTCTTGTTCTTCTTATATTTCTCTCTCTGTAAGT

TTTCCCCTCTCTCTCTTTCTCTCTCCATAAGAACTCACCAAGAACACATAGAACACGACCAAGAACAATAATCTTAGATTTTT

TTTATCTGCACCCACTTCTTATCTTTTACAATAATTTTTGAGACAGTGTTAATGATTGACAGTGTTAAGTTGGAATATTTTTA

TTTTTTTGAAACTGAATAATTTGTGAGACCAATCTCATTTAAAGTATATGTAACATTTAATTGAGTTCAATATTTATTTAATT

GTATTAATGGATTTGCTTTATATGTGGACAGTGATTATCCGCCGTTTCTGCTTAACTGTGGATTGTTGTTTCTTTGATCTTTG

AAGAATTCAGTC**A**

2000 bp

Appendix 4 – Tables and figures for chapter 7

Appendix A4.1. Normalized expression value for *AtCDK8*, *AtCYCC1-1* and *AtCYCC1-2* genes obtained after various durations from control, heat, UVB, wound and salt treated plants.

Tissue cluster	Tissue	Time	Treatment	Normalized expression value		
				AtCDK8	*AtCYCC1-1*	*AtCYCC1-2*
Control, aerial	Seedling, green parts	0.5 h	Control: The plants were treated like the real treated plants, that means: Get boxes out of the climate chamber. Open the magenta box and lift the raft as long as the real treatment lasts. Then put them back in the climate chamber. Harvested 0.5-24 h after start of experiment	0.9259	0.848	1.6261
		1 h		0.7408	0.971	1.5553
		3 h		0.7563	0.9761	1.4945
		6 h		0.6606	1.0311	1.6168
		12 h		0.6899	1.4199	1.4866
		24 h		0.8382	1.0591	1.9355
Heat, aerial		0.5 h	0.5- 3 h of 38° C heat stress in an incubator	0.7157	0.9529	1.6286
		1 h		0.6545	0.902	1.7214
		3 h		1.0195	1.0151	1.1981
		6 (3hr + 3h recovery) h	3.0h of 38° C heat stress in an incubator followed by 3-12 h recovery at 25° C	0.7427	1.2027	1.8107
		12 (3h + 9h recovery) h		0.693	1.6227	1.986
		24 (3h + 21h recovery) h		0.8137	1.0418	2.1345
UV-B, aerial		0.5 h	UV-B stress: The plants were stressed by 0,25 h UV-B light field consisting of six Philips TL 40W/12 UV fluorescent tubes (_max. 310 nm, half-bandwidth 40 nm, fluence rate 7 W/m^2) filtered through 3 mm transmission cutoff filters of the WG series ((WG295, WG305, and WG327; Schott, Mainz, Germany). Harvested 0.5-24 h after start of experiment.	0.6763	1.0203	1.6473
		1 h		0.6876	1.0293	1.7002
		3 h		0.8788	1.0457	1.0394
		6 h		0.8471	0.915	0.9163
		12 h		0.7218	1.3417	1.6884
		24 h		0.7229	1.1751	1.561
Wound, aerial		0.5 h	Wound stress: The plants were wounded by punctuation of the leaves with a custom made pin-tool consisting of 16 needles (~2 needles/1 cm^2). Three times of consecutive application pierced on average three to four distinct holes per leaf. Harvested 0.5 - 24 h after start of experiment	0.7086	0.8176	1.6825
		1 h		1.1155	0.7979	1.3492
		3 h		0.7601	1.067	1.3918
		6 h		0.8824	1.1795	1.4319
		12 h		0.7342	1.3259	1.7184
		24 h		0.9172	1.1303	1.7372
Salt, aerial		0.5 h	Salt stress (150 mM NaCl): NaCl was added to a concentration of 150 mM in the Media. To add the NaCl the raft was lifted out. A magnetic stir bar and a stirrer was used to mix the media and the added NaCl. Then the raft was put back in the box. The boxes were put back to the climate camber. Harvested 0.5 -24 h after start of experiment	1.078	0.8525	1.5324
		1 h		0.7351	1.0531	1.4987
		3 h		0.7457	0.8608	1.0396
		6 h		0.7739	0.8127	1.0261
		12 h		0.7949	0.9188	1.2633
		24 h		0.7759	0.8596	1.0278

* Normalized values are obtained by normalizing absolute values (linearized gcRMA values) to median for each gene across all samples. Robust Multi-Array Average (RMA) is a normalisation procedure used in microarrays that corrects, normalises and summarises the background probe level information whereas GeneChip-RMA (GC-RMA) is an improved form of RMA that exploits the sequence-specific probe affinities of the GC probes delivering more precise gene expression values.

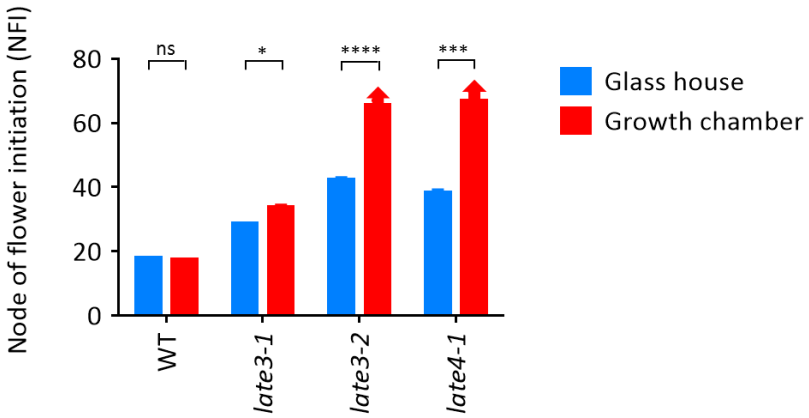


Figure A4.1. Responsiveness of WT (NGB5839), *late3* and *late4* mutants to light conditions under LD, 20° C in two different places. Glasshouse contains both natural and artificial light whereas growth chamber has only artificial light. *late3-2* and *late4-1* mutants did not flower in growth chamber and graphs show the total number of nodes at which these mutants underwent senescence; for all other cases in each the genotypes the graph is showing the node at which flowering initiation occurred. Data represents mean ± SE for n =4-6 plants. Data from figure 3.1 have been re-plotted. Welch t-test (two tailed) was performed to determine the level of significance using Prism software. p-value significance level P ≤ 0.0001, ***≤ 0.001, ** ≤0.01, *≤ 0.05, ns > 0.05.